

NASA CR-165, 111

NASA-CR-165111
19820008228

The Telecommunications and Data Acquisition Progress Report 42-66

September and October 1981

N. A. Renzetti
Editor

December 15, 1981



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

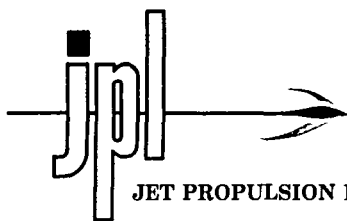
LIBRARY COPY

JAN 21 1982

LANGLEY RESEARCH CENTER
LIBRARY, NASA
HAMPTON, VIRGINIA



NF02423



JET PROPULSION LABORATORY California Institute of Technology • 4800 Oak Grove Drive, Pasadena, California 91103

February 15, 1982

Recipients of Jet Propulsion Laboratory
Tracking and Data Acquisition Progress Report 42-66

SUBJECT: Errata

Please note the following corrections to the article entitled "An M-ary Coherent Optical Receiver for the Free-Space Channel," by V.A. Vilmrotter, TDA Progress Report 42-66, Dec. 15, 1981, pp. 60-66.

An error was made in expanding Eq. (11), which led to an error in Eq. (12b). Since Eq. (12a) is the limiting form of Eq. (9), and since Eq. (11) is contained in Eq. (9), (12a) was also affected by the same error. The correct limits are:

$$\Psi_{11}(\omega) \xrightarrow{\lambda_L \rightarrow \infty} j\omega(2K_S) - \frac{\omega^2}{2}(4K_S) \quad (12a)$$

$$\Psi_{11}(\omega) \xrightarrow{\lambda_L \rightarrow \infty} j\omega(-2K_S) - \frac{\omega^2}{2}(4K_S) \quad (12b)$$

with the interpretation that (12a) is the second characteristic function of a Gaussian random variable with mean value $(2K_S)$ and variance $(4K_S)$, while (12b) corresponds to a Gaussian random variable with mean value $(-2K_S)$ and variance $(4K_S)$. Since this error merely introduced a shift in the mean, it has no effect on any of the subsequent equations or results, all of which remain valid.

A clarification is also in order: Although not stated explicitly, it can be shown that in the limit as λ_L approaches infinity, the $\{\Lambda_i\}$ become statistically independent Gaussian random variables; hence Eq. (13) follows.

Very truly yours,

John Kempton, Manager
Publications Section

The Telecommunications and Data Acquisition Progress Report 42-66

September and October 1981

N. A. Renzetti
Editor

December 15, 1981



**National Aeronautics and
Space Administration**

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

N82-16101 #

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration

Preface

This publication was formerly entitled *The Deep Space Network Progress Report*. Although the practice of reporting progress in the development and operations of the Deep Space Network continues, the report has been expanded to include developments in Earth-based radio technology as applied to other research programs. These programs are:

- (1) **Geodynamics** For several years, the laboratory has been developing radio interferometry at microwave frequencies for application to geodetic measurements. This branch of telecommunications technology is now being applied to the study of geodynamics.
- (2) **Astrophysics** The deep space stations, individually and in pairs as an interferometer, have been used by radio astronomers for astrophysics research by direct observations of radio sources.
- (3) **An activity closely related to radio astronomy's use of the deep space stations is NASA's continuing program of radio search for extraterrestrial intelligence in the microwave region of the electromagnetic spectrum.**

Each succeeding issue of this report will present material in some, but not all, of the following categories:

- Radio Astronomy
- Search for Extraterrestrial Intelligence
- Radio Interferometry at Microwave Frequencies

- Geodetic Techniques Development
 - Spacecraft Navigation
 - Orbiting Very Long Baseline Interferometry

- Deep Space Network

- Description
 - Program Planning
 - Planetary and Interplanetary Mission Support
 - Advanced Systems
 - Network and Facility Engineering and Implementation
 - Operations
 - Spacecraft Radio Science
 - Planetary Radar
 - Energy

In each issue, there will be a report on the current configuration of one of the seven DSN systems (Tracking, Telemetry, Command, Monitor and Control, Test Support, Radio Science, and Very Long Baseline Interferometry).

The work described in this report series is either performed or managed by the Telecommunications and Data Acquisition organization of JPL.

This Page Intentionally Left Blank

Contents

RADIO ASTRONOMY

Radio Astronomy	1
R D Shaffer, G S Levy, H C Wilck, M J Grimm, S Gulkis, T B Kuiper, and E T Olsen NASA Code 311-03-21-00	

RADIO INTERFEROMETRY AT MICROWAVE FREQUENCIES GEODETIC TECHNIQUES DEVELOPMENT

Proper Frequencies of a Down-Looking Water-Vapor Radiometer Over Sea Surface	4
S C Wu NASA Code 692-40-01-00	

SPACECRAFT NAVIGATION Block I VLBI

Block I Phase 2 Very Long Baseline Interferometry Implementation	15
J H Wilcher NASA Code 311-03-41-01	

THE DEEP SPACE NETWORK DESCRIPTION OF THE DSN

The DSN Programming System	19
G A Madrid NASA Code 311-03-41-01	

PLANETARY AND INTERPLANETARY MISSION SUPPORT Planetary Flight Projects

Voyager Mission Support (I)	24
N Fanelli and H Nance NASA Code 311-03-22-20	
Voyager Mission Support (II)	27
N Fanelli and H Nance NASA Code 311-03-22-02	

ADVANCED SYSTEMS Tracking and Ground-Based Navigation

A VLBI Survey of the Northern Polar Cap Region	30
D D Morabito, R A Preston, and J Faulkner NASA Code 310-10-60-26	
Arcsecond Positions for Milliarcsecond VLBI Nuclei of Extragalactic Radio Sources. Part I 546 Sources	37
D D Morabito, R A Preston, and M A Slade NASA Code 310-10-60-26	

An FM/CW Method for the Measurements of Time Delays of Large Cassegrain Antennas	49
T Y Otsu	
NASA Code 310-10-60-05	

Communications

An M-ary Coherent Optical Receiver for the Free-Space Channel	60
V A Vilnrotter	
NASA Code 310-40-73-55	
Thermal Measurement Technique of Rib Elements on DSN Antenna Structure	67
S Glazer and G Gale	
NASA Code 310-20-65-04	
X-Band Narrow-Beam Radiometer for DSS 13	80
M M Franco, S D Slobin, and C T Stelzried	
NASA Code 310-20-66-15	
Design of a 1.5-m, 32-GHz, Clear Aperture Antenna	87
A Cha	
NASA Code 310-20-65-04	
Power Efficiency of Semiconductor Injection Lasers	94
J Katz	
NASA Code 310-20-67-59	
Phase-Locking of Semiconductor Injection Lasers	101
J Katz	
NASA Code 310-10-67-59	
2.5-Bit/Detected Photon Demonstration Program Description, Analysis, and Phase I Results	115
J R Lesh, J Katz, H H Tan, and D Zwillinger	
NASA Code 310-20-67-59	

Station Control and System Technology

DSS 13 Unattended Operations Station Controller Status Report	133
H Cooper	
NASA Code 310-20-67-54	
A Digital Filter Implementation of the Deep Space Transponder Phase Lock Loop Integrator	138
J W Heller	
NASA Code 310-30-70-00	

Network Data Processing and Productivity

Improved Carrier Tracking Performance with Coupled Phase-Locked Loops	148
D Divsalar and J H Yuen	
NASA Code 310-40-74-57	
Implications of the Putnam Software Equation on Confidence Levels for Project Success	172
R C Tausworthe	
NASA Code 310-40-72-10	
Gain Uncertainty of Arrayed Antennas	186
B L Seidel	
NASA Code 310-40-74-54	

Life-Cycle Cost Analysis of Projects Using a Polynomial Cash Flow Model for Nonuniform Maintenance and Operations Costs . . .	191
D S Remer and G Lorden	
NASA Code 310-40-73-06	

NETWORK AND FACILITY ENGINEERING AND IMPLEMENTATION

Network

Sine Wave Ranging Revisited .	202
A I Zygielbaum and A Robinson	
NASA Code 311-03-41-91	
Radio Frequency Interference by Earth Orbiting Satellites: Deep Space Interference Prediction Program .	211
D L Cain and P E Beyer	
NASA Code 311-03-41-31	
New S-Band Transmitter Automation Software .	214
W Stahnke	
NASA Code 312-03-42-09	
The Effect of Partial Coherence in Receiving System Noise Temperature on Array Gain for Telemetry and Radio Frequency Carrier Reception for Similar Receiving Systems	219
M H Brockman	
NASA Code 311-03-43-20	
Optimum Equipment Maintenance/Replacement Policy Part 1: Dynamic Programming Approach	236
T Charng	
NASA Code 311-03-41-08	
System Reliability Analysis with Dynamic Programming .	246
T Charng	
NASA Code 311-03-41-08	
A Computerized Life-Cycle Cost Methodology for Engineering Analysis .	268
R D Hughes	
NASA Code 311-03-41-08	

Ground Communications

GCF-NOCC Reconfiguration	288
J P McClure	
NASA Code 311-06-20-00	

Deep Space Stations

Antenna Azimuth Bearing Model Experiment	291
H McGinness	
NASA Code 311-20-65-05	

OPERATIONS

Network Operations

Operational VLBI Clock Synchronization and Platform Parameter Determination	307
J V LuValle, R D Shaffer, M G Roth, T M Eubanks, and P S Callahan	
NASA Code 311-03-141-296	

Synthesizer Stability Evaluation . . . 319
 E N Sosa and D A Tyner
 NASA Code 312-03-51-32

Investigation of Digitally Controlled Oscillator Operability Problems . 331
 W D Weisman
 NASA Code 311-03-12-10

Emergency Telemetry Performance . . 335
 W Tisdale
 NASA Code 311-03-14-20

Deep Space Stations

Goldstone STDN 9-Meter Radiation Test . . . 341
 J R Blain
 NASA Code 311-03-14-20

ENERGY

A Solar Powered Water Purification System at Goldstone (DSS 13) . 352
 F J Menninger and R J Elder
 NASA Code AA-882-20-00-00

Energy Consumption for the Echo Station (DSS 12) . 355
 C N Guiar and D Schonfeld
 NASA Code 311-03-41-08

Radio Astronomy

R. D. Shaffer
Control Center Operations Section

G. S. Levy
TDA Technology Development Section

H. C. Wilck and M. J. Grimm
Communications Systems Research Section

S. Gulkis, T. B. Kuiper, and E. T. Olsen
Planetary Atmospheres Section

This article reports on the activities of the Deep Space Network in support of Radio Astronomy Operations during September and October 1981. Work in progress in support of a program sponsored by the Director's Discretionary Fund is reported.

I. Introduction

Deep Space Network (DSN) 26-, 34- and 64-meter antenna stations are utilized in support of three categories: NASA Office of Space Science (OSS), Radio Astronomy Experiment Selection (RAES), and Host Country. This report also features progress of the Science Utilization of the Radio Frequency Interference Surveillance System (SURFISS). While this activity does not fit the categories on which we normally report, it is basically radio astronomy, and thus is included here.

II. Radio Astronomy Operations

A. NASA OSS Category

During this period, 73 hours and 35 minutes of DSN station support was given for Pulsar Rotation Constancy, 17 hours and 30 minutes of support was given for Planetary Radio Astronomy, and 8 hours and 15 minutes of support was given for Southern Hemisphere Radio Source positions.

B. RAES Panel Category

While no primary DSN support for any RAES experiment was rendered during this reporting period, RA 178 (superluminal radio sources at 13 cm) was approved by the RAES panel and scheduled for DSN support in late 1981. Another new proposal, "A Statistical VLBI Test of the Self-Compton Mechanism for X-Ray Emission from Compact Extragalactic Radio Sources," was also received and is being reviewed by the RAES panel members.

C. Host Country

At time of going to press, the figures for this activity were unavailable.

III. SURFISS Progress Report

Science Utilization of the Radio Frequency Interference Surveillance System (SURFISS), a program sponsored by the

Director's Discretionary Fund, is designed to utilize JPL's 65,000-channel Radio Frequency Interference Surveillance System (RFISS) for radio astronomical studies. The RFISS is a unique microwave spectrum analyzer developed to detect radio frequency interference (RFI) at the DSN facility at Goldstone. It is a trailer-based microwave receiver and data processing system which includes a 65,000-channel spectrum analyzer. The receiver and spectrum analyzer are controlled by a Modcomp II computer, allowing the RFISS to process two 10-MHz channels into 32,000 channels, each with a resolution of 305 Hz.

The RFISS was connected to the 26-m antenna receiver system at DSS 13 (Venus Station) in September 1981 in order to carry out the first step of the SURFISS program. Both S- and X-band signals were coupled to the RFISS for these tests. The objective of the series of tests which took place on September 10/11, 17/18, and 24/25 was to obtain experience with the RFISS as the processor for DSS 13 output and to investigate the system stability and gain characteristics. Observational procedures were tested for first-order calibration of data.

The data obtained during the five observing runs are analyzed both in the field and off-line on a PDP 11/44 at JPL. The software available on the Modcomp II in the field allows the observer to inspect spectra in near-real-time in both normalized and calibrated format. Thus the observer may make informed decisions in real-time. For off-line analyses, the data are written on magnetic tape.

Astronomical observations were carried out October 8/9. During this run, recombination lines in M17, Orion A, and W51 at S-band were observed. $H142\alpha$ and $He142\alpha$ were seen in all three sources of thermal microwave emission. $H178\alpha$ was seen in M17 and Orion A. At X-band, the transition $H115\beta$ was seen in Orion A, and an unidentified line was seen in M17. Figure 1 is a calibrated spectrum of M17 at S-band showing the $H142\alpha$ and $He142\alpha$ recombination lines. Since these lines are several hundred kilohertz wide, the frequency resolution of the data has been degraded through averaging adjacent channels from the original 305 Hz to 16.16 kHz. The vertical scale is antenna temperature in kelvin, the horizontal scale is frequency in MHz. Figure 2 is a similarly calibrated spectrum of Orion A at X-band showing the $H115\beta$ recombination line.

Observations of Voyager 2 at X-band and Pioneer 12 at S-band were also made. Figure 3 is a calibrated spectrum of Voyager 2 showing the suppressed carrier and upper- and lower-sidebands out to the ninth sideband, at a resolution of 16.16 kHz. Figures 4 and 5 are calibrated spectra of the first lower- and upper-sidebands, and illustrate the asymmetry which is present in the signal.

Work is currently underway to evaluate the performance of the system when observing astronomical sources and spacecraft. It is planned to conduct observations at DSS 14 in January 1982 after the completion of a series of RFI surveys at both DSS 13 and DSS 14.

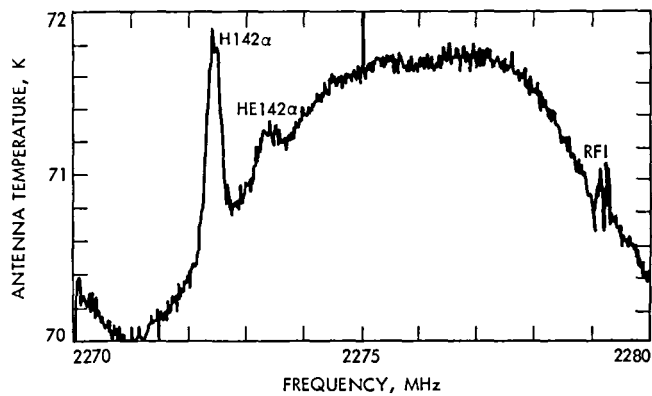


Fig 1. M17 at S-band

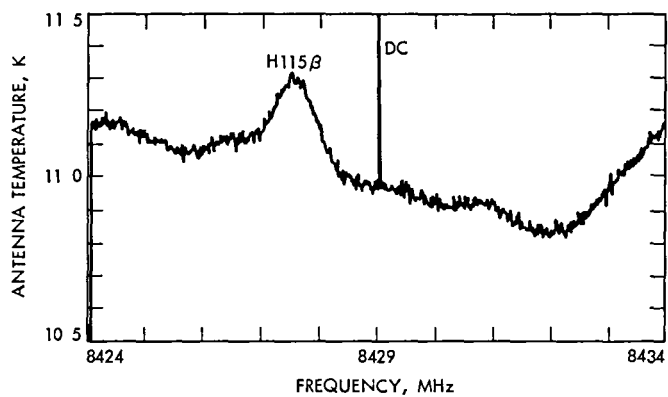


Fig. 2. Orion A at X-band

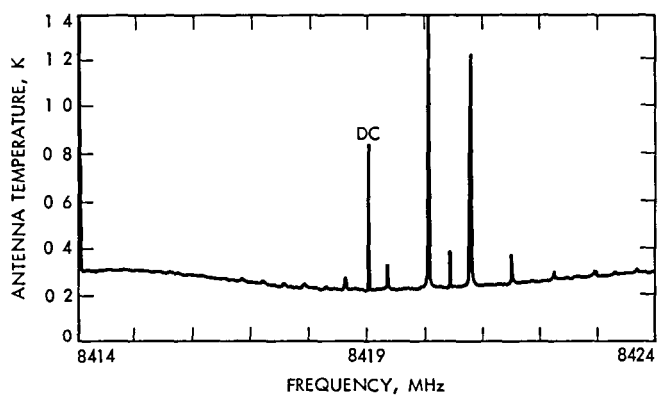


Fig 3. Voyager 2 calibrated spectrum

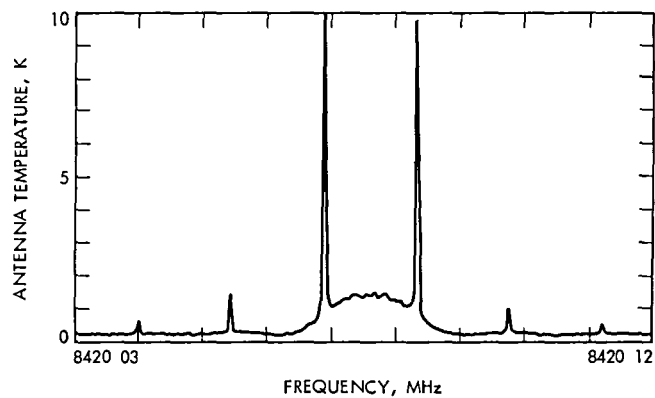


Fig 4. Voyager 2, lower 1st sideband

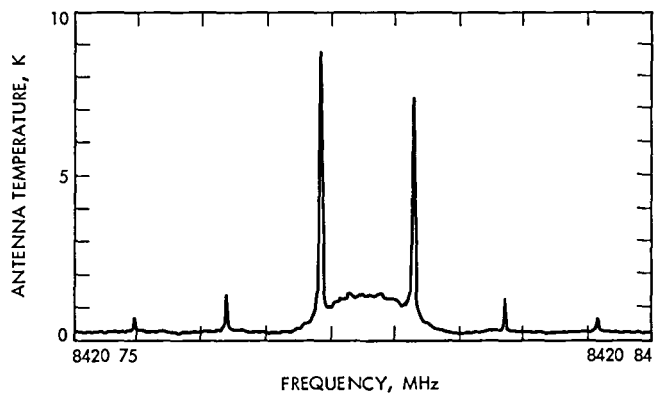


Fig. 5 Voyager 2, upper 1st sideband

Proper Frequencies of a Down-Looking Water-Vapor Radiometer Over Sea Surface

S. C. Wu

Tracking Systems and Applications Section

Altimeter observation on board an Earth-orbiting satellite (e.g., TOPEX) over sea surface is corrupted by, among others, the water vapor path delay. Such error can be calibrated for by a down-looking microwave radiometer. This article studies the effects of water vapor profile, sea surface temperature, wind speed and brightness temperature measurement error on the calibration precision. Proper frequency combinations are searched to minimize these effects.

I. Introduction

Tropospheric water vapor introduces path delay as large as 50 cm on radio propagation. The most effective calibration method has been that making use of passive microwave radiometers (Refs. 1-5). The thermal radiation of water vapor at frequencies near the 22.235-GHz water absorption line, which can be measured by a radiometer, is an accurate indication of water vapor content along the line of sight.

On a clear day when no cloud (liquid water droplets) is present, a ground-based single-frequency measurement of the sky brightness temperature is sufficient for the estimation of the water vapor content or the corresponding delay. When the line of sight passes through clouds, the observed brightness temperature increases rapidly even though the path delay is hardly affected. The cloud effect can be separated from the water vapor effect by making measurement at two different frequencies. Such separation is possible owing to the difference in radiation frequency spectra between water vapor and liquid water droplets.

Another factor affecting the accuracy of water vapor measurement is the variation of the water vapor vertical profile. This is due to the fact that the thermal radiation from an amount of tropospheric constituents (including water vapor) farther away from the radiometer is continuously attenuated (through absorption) by the nearer constituents along the line of sight, the same amount of water vapor at a different height will contribute to different brightness temperature at the radiometer. Such profile dependence can be alleviated by the selection of a proper frequency (Refs. 2, 6) or a combination of frequencies (Refs. 3, 7) for the radiometer.

When the measurement is made on board an aircraft or an Earth-orbiting satellite over a sea surface, the problem becomes more complicated than that of a ground-based measurement in two respects. First, the brightness temperature is now the combination of three components: the upward tropospheric radiation, the reflected downward tropospheric radiation and the sea surface radiation. Secondly, both the reflected and the radiated components by the sea surface are strongly dependent

on sea surface emissivity which in turn is a function of sea surface temperature and foam coverage. To achieve an accurate estimation of water vapor path delay, the variations of sea surface temperature and emissivity cannot be ignored.

In this article, we shall investigate whether sub-centimeter water-vapor delay is achievable using a down-looking radiometer with proper selection of two frequencies. The effects of vertical profile, sea surface temperature and wind speed (related to foam coverage), and brightness temperature measurement error will be studied. The absence of precipitation will be assumed.

II. Background

The quantity to be determined is the water vapor path delay (Ref. 8)

$$\Delta R = k \int_0^H \frac{\rho_v}{T} dh, \quad k = 1.723 \times 10^{-3} \text{ K/(g/m}^3\text{)} \quad (1)$$

where ρ_v is the water vapor density and T is the atmospheric temperature along the vertical ray path h , H is the altitude of the radiometer. The quantity which is measured by a radiometer is the brightness temperature

$$\begin{aligned} T_B = & \int_0^H T \alpha \exp \left(- \int_h^H \alpha dh \right) dh \\ & + e^{-\tau} \left[\epsilon_s T_s + (1 - \epsilon_s) \left(\int_0^H T \alpha \exp \left(- \int_0^h \alpha dh \right) dh \right. \right. \\ & \left. \left. + T_c e^{-\tau} \right) \right] \quad (2) \end{aligned}$$

Here, the first term is due to the upward tropospheric radiation, the terms in the brackets are due to the sea surface radiation and the sea surface reflections of the downward tropospheric radiation and of the cosmic radiation. These bracketed terms are attenuated by the tropospheric absorption $\tau = \int_0^H \alpha dh$, with $\alpha = \alpha_v + \alpha_c + \alpha_o$ being the combined absorption coefficient of water vapor, cloud and oxygen along the ray path. The sea surface emissivity is denoted by ϵ_s and the sea surface temperature by T_s . The cosmic radiation has a value of $T_c = 2.9 \text{ K}$. From Eq. (2) the dependence of T_B on ϵ_s , T_s and the vertical profile of α may be appreciated.

To measure the water vapor path delay by a radiometer a calibration equation relating the two quantities, ΔR in (1) and T_B in (2), needs to be established.

III. Calibration Equation

The brightness temperature of Eq. (2) can be expressed as the following approximation (Ref. 9. In Ref. 9 the cosmic radiation T_c has been omitted)

$$T_B \doteq T_e + e^{-\tau} \epsilon_s (T_s - T_e) - e^{-2\tau} (1 - \epsilon_s) (T_e - T_c) \quad (3)$$

where T_e is the effective temperature defined as

$$T_e = \frac{\int_0^H T \alpha \exp \left(- \int_h^H \alpha dh \right) dh}{\int_0^H \alpha \exp \left(- \int_h^H \alpha dh \right) dh} \quad (4)$$

The absorption τ , which is the algebraic sum of the contributions from water vapor, cloud and oxygen, increases *linearly* with the amount of water vapor content along the ray path. Hence it is the logical quantity to be related to ΔR . Solving for τ from (3) we have

$$\tau = \ln \left[\frac{2(1 - \epsilon_s)(T_e - T_c)}{\epsilon_s(T_s - T_e) + \sqrt{\epsilon_s^2(T_s - T_e)^2 + 4(1 - \epsilon_s)(T_e - T_c)(T_e - T_B)}} \right] \quad (5)$$

In the calculation of τ using (5), the effective temperature T_e needs to be accurately estimated. To loosen the dependence on the estimation of T_e , a quantity $T_e^\gamma \tau$ is to be used instead of τ . The exponent γ is selected such that $T_e^\gamma \tau$ is insensitive to the variation of T_e . In other words, $\partial(T_e^\gamma \tau)/\partial T_e = 0$. The derivative is rather lengthy but straightforward. After calculating the derivative and using the typical values $T_s = 290 \text{ K}$, $T_e = 275 \text{ K}$, $T_B = 150 \text{ K}$ and $\epsilon_s = 0.45$ we have $\gamma \doteq 2$. Therefore, $T_e^2 \tau$ is to be used as the parameter in the calibration equation.

It can be numerically shown that

$$T_e^2 \tau \doteq \int_0^H T^2 \alpha dh$$

or

$$T_e^2 \tau \doteq \int_0^H T^2 (\alpha_v + \alpha_c + \alpha_o) dh \quad (6)$$

Since α_c is proportional to f^2 (Ref. 10) for microwave frequencies $f < 40$ GHz, it can be eliminated for a dual-frequency radiometer

$$\frac{(T_e^2 \tau)_1}{f_1^2} - \frac{(T_e^2 \tau)_2}{f_2^2} = \int_0^H w(h) \frac{\rho_v}{T} dh + T_{ox}^2 \quad (7)$$

where

$$w(h) = \frac{T^3}{\rho_v} \left(\frac{\alpha_{v,1}}{f_1^2} - \frac{\alpha_{v,2}}{f_2^2} \right) \quad (8)$$

and

$$T_{ox}^2 = \int_0^H T^2 \left(\frac{\alpha_{o,1}}{f_1^2} - \frac{\alpha_{o,2}}{f_2^2} \right) dh \quad (9)$$

A comparison of (1) and (7) indicates that a linear relationship between ΔR and $T_e^2 \tau$ at the two frequencies can be established provided the weighting function $w(h)$ be a constant, independent of h . Under such circumstances,

$$\Delta R = a_0 + a_1 (T_e^2 \tau)_1 + a_2 (T_e^2 \tau)_2 \quad (10)$$

where the constant coefficients are

$$\begin{aligned} a_0 &= \frac{-k}{w} T_{ox}^2 \\ a_1 &= \frac{k}{w f_1^2} \\ a_2 &= \frac{-k}{w f_2^2} \end{aligned} \quad (11)$$

For $f < 40$ GHz, the oxygen absorption coefficient α_o also increases with approximately f^2 . Hence T_{ox}^2 as defined by (9) is a small residual oxygen effect. Furthermore, α_o is proportional to $T^{-2.85}$ (Ref. 11), $T^2 \alpha_o$ is insensitive to the temperature T . Therefore, T_{ox}^2 is practically a constant. Note that a constant $w(h)$ implies the estimation of ΔR being independent of the water vapor vertical profile. In the following section, frequency pairs for which $w(h)$ remains nearly constant will be searched.

IV. Frequency Pairs for Constant $w(h)$

The procedure of Ref. 7 is to be followed in the search for frequency pairs resulting in nearly constant $w(h)$. In brief, the

single frequency contributions of $w(h)$ in (8) are calculated and plotted on the same graph, as depicted by Fig. 1, any pair (e.g., b and c in Fig. 1) having a wide and nearly constant separation will result in a large and constant $w(h)$ and can be selected as a candidate for further consideration. The actual calculation will be based on the radiosonde measurements of the meteorology profiles at Point Mugu, California, on February 24, 1976. These profiles are shown in Fig. 2. It has been shown in Ref. 7 that, for a ground-based radiometer with specified frequencies, the shape of $w(h)$ vs h is similar for different profile shapes. Since the form of $w(h)$ in (8) differs only by a multiplicative factor T which normally has a small fractional variation, the same can be said for a down-looking radiometer. Hence the profile of Fig. 2 and the resulting $w(h)$ can be considered as typical. Note that we need to examine $w(h)$ only for h below about 5 km since there is little water vapor content at higher altitudes.

Figure 3 shows the components of $w(h)$ for frequencies near the 22.235-GHz water absorption line. Two of the pairs having wide and nearly constant separations are found to be 20.2 GHz/26.0 GHz and 24.3 GHz/32.0 GHz. The normalized variations in $w(h)$ of these frequency pairs are shown in Fig. 4. Also shown are the 20.3 GHz/31.4 GHz optimum pair for a ground-based radiometer (Ref. 7), the 22.235 GHz/31.4 GHz pair selected by the Nimbus-6 Scanning Microwave Spectrometer (SCAMS) (Ref. 9) and the pair 18.0 GHz/21.0 GHz believed to be insensitive to sea surface temperature¹. Note that the last two pairs have large $w(h)$ variations. The pair for the optimum ground-based radiometer (20.3 GHz/31.4 GHz) is only slightly inferior to the remaining two at higher altitudes. The pair with least variation in $w(h)$ is found to be 24.3 GHz/32.0 GHz.

The values of T_{ox}^2 in (9) and of $w(h)$ in (8) at sea surface ($h = 0$) are calculated and shown in Table 1. The 24.3 GHz/32.0 GHz pair not only has a reasonably large $w(0)$ for small sensitivity to $T_e^2 \tau$ error when (10) is used to estimate the water vapor path delay, it also has the lowest T_{ox}^2 , which implies being least affected by the variation in the residual oxygen term.

V. Effects of Sea Surface Temperature and Wind Speed

The variation in sea surface temperature affects the estimation of the water vapor path delay in two respects. First, it varies the values of the weighting function w , which in turn varies the calibration equation coefficients (cf. Eq. 11). Secondly, it varies the observable $T_e^2 \tau$ through T_s and ϵ_s (cf.

¹T. Chester, private communication

Eq. 5) The variation in sea surface wind speed has an effect only through ϵ_s .

To study the effect of sea surface temperature and wind speed through ϵ_s , it is convenient to work out an empirical mathematical expression for ϵ_s in terms of f , T_s and V_w , the sea surface wind speed. Figure 5 shows the frequency dependence of ϵ_s at $T = 293$ K (Ref. 12). The effect of sea water salinity is found to be negligibly small for $f > 5$ GHz. The area which concerns us is the nadir ray path in the frequency range of 15 to 35 GHz. Within this range ϵ_s is nearly linear in frequency and can be represented by

$$\epsilon \doteq 0.34(1 + 0.01f), \quad T = 293 \text{ K} \quad (12)$$

The temperature dependence of ϵ_s can be derived from Fig. 6 (Ref. 12), where the change in $T_s \epsilon_s$ per change in T_s as a function of f is shown. This figure applies to the temperature range of 273 to 303 K, practically all possible sea surface temperatures. Over the frequency range of interest, a close approximate expression for the nadir observation is

$$\frac{\Delta(T_s \epsilon_s)}{\Delta T_s} \doteq 1.46 e^{-0.04f} - 0.64 \quad (13)$$

Wind over the sea surface increases the emissivity through the amount of foam coverage. Nordberg et al. (Ref. 13) found that, at $f = 19.35$ GHz, there is no detectable effect for wind speeds less than 7 m/sec while $T_s \epsilon_s$ increases by 1 K per m/sec above that, as shown in Fig. 7. The effect is nearly frequency-independent, as has been discovered by Webster et al. (Ref. 14), in the frequency range we are concerned with. This is shown in Fig. 8. The effect of wind speed can be expressed as

$$\frac{\partial(T_s \epsilon_s)}{\partial V_w} \doteq \delta_{-1}(V_w - 7) \quad (14)$$

where $\delta_{-1}(x)$ is the unit step function defined by

$$\delta_{-1}(x) = \begin{cases} 0 & , \quad x < 0 \\ 1 & , \quad x \geq 0 \end{cases}$$

Combining (12), (13) and (14) we arrive at the following empirical expression for the sea surface emissivity.

$$\begin{aligned} \epsilon_s = & \frac{99}{T_s} (1 + 0.01f) + \left(\frac{T_s - 293}{T_s} \right) (1.46 e^{-0.04f} - 0.64) \\ & + \delta_{-2}(V_w - 7)/T_s \end{aligned} \quad (15)$$

where $\delta_{-2}(x)$ is the unit ramp function defined by

$$\delta_{-2}(x) = \begin{cases} 0 & , \quad x < 0 \\ x & , \quad x \geq 0 \end{cases} \quad (16)$$

The sea surface value of the weighting function, $w(0)$, varies with the surface values of T^3 and α_v/ρ_v (cf. Eq. 8). The variation in α_v/ρ_v is small as compared to that in T^3 . Hence it is a good approximation to assume $w(0)$ as being dependent only on T in the form

$$w(0) \propto T_s^3 \quad (17)$$

VI. Numerical Results

In this section, we shall study the statistical uncertainty in the estimation of water vapor path delay using the calibration equation (10) with different frequency pairs. Five sets of theoretical coefficients are calculated from the values of T_{ox}^2 and $w(0)$ in Table 1 of the typical meteorology profiles in Fig. 2.

The test data are a set of 75 meteorology soundings derived from expanding 25 measurements at Point Mugu, California, from February 24 to March 17, 1976. The expansion is furnished by simply adding -10 K and 10 K to the temperature profiles, to simulate the worldwide coverage of temperature range. The brightness temperatures that would be measured by a radiometer are calculated by (2). The measurements cover all altitudes up to ~10 km, above which the water vapor content is rare. These data are extrapolated to an altitude at which the total pressure drops to 100 mbar ($H \doteq 16$ km) to include ~99% of the oxygen contribution to T_B .

The theoretical coefficients derived from Table 1 and the brightness temperatures calculated from the 75 soundings are used to estimate the water vapor path delay by Eq. (10). The effective temperature T_e is found to be ~0.95 T_s for all frequencies. The estimated ΔR are to be compared with the actual ΔR calculated from (1). The statistics of their differences provide an estimate of the calibration equation uncertainty.

Four different test cases are to be studied

- Case 1. No wind ($V_w < 7$ m/sec)
Coefficients scaled by T_s^{-3}
 τ calculated by correct T_s
- Case 2. No wind
Coefficients fixed
 τ calculated by fixed T_s (287.6 K from Fig. 2)

- Case 3. 10 ± 5 m/sec wind
Coefficients scaled by T_s^{-3}
 τ calculated by $V_w = 10$ m/sec and correct T_s
- Case 4. 10 ± 5 m/sec wind
Coefficients fixed
 τ calculated by $V_w = 10$ m/sec and $T_s = 287.6$ K

Table 2 summarizes the results of the comparison. The following points are observed

- (1) When independent measurements of sea surface temperature and wind speed are available and are applied to the calculations of τ and the calibration coefficients (Case 1), the uncertainties are less than 1 cm for all five frequency pairs.² The three having nearly constant $w(h)$ in Fig. 4 are 3-5 times better than the remaining two. This verifies the reduction of profile dependence of Eq. (10) with proper frequency pairs.
- (2) When sea surface temperatures (having a standard deviation $\sigma_{T_s} = 9.2$ K) are not independently measured (Case 2), the uncertainties increase by an appreciable amount for the first four frequency pairs. The pair 18 GHz/21 GHz is nearly unchanged. Such insensitivity to sea surface temperature is a result of compensation between the change in w and the change in $T_e^2 \tau$ due to a change in T_s .
- (3) When the sea surface wind speed is mis modeled by $\sigma_{V_w} = 5$ m/sec (Case 3), the uncertainties in ΔR increase by 0.8 - 1.4 cm. The pair 20.3 GHz/31.4 GHz has the largest increase and the pair 22.235 GHz/31.4 GHz has the least.
- (4) When the mis modeling includes both $\sigma_{V_w} = 5$ m/sec and $\sigma_{T_s} = 9.2$ K (Case 4), the combined effects are nearly the RSS of the two separate effects. The frequency pair having the most constant $w(h)$ - 24.3 GHz/32.0 GHz - is still better than the other four, with an uncertainty of ~ 1.5 cm.

The uncertainties above have been estimated using theoretical coefficients in the calibration equation (10). These uncertainties would have been lower had the coefficients been derived from regression analysis with the 75 soundings. For comparison, these estimates are summarized in Table 3. It is believed that such "forced-fitting" may result in an optimistic

²The large biases for the last two frequency pairs are due to an underestimate of w by the surface values. Such biases can be reduced by using w at an altitude above sea surface for the calculation of coefficients in (11)

estimate. An ideal error estimation approach would be such that the calibration coefficients are derived from a regression analysis with one set of data and are applied to another set of data for ΔR error estimation

Another error source not included in the above analysis is in the brightness temperature measurement. The sensitivity to such error can be estimated by omitting the term containing $e^{-\tau}$ in (3) and then substituting into (10) and taking partial derivative with respect to T_B .

$$\frac{\partial(\Delta R)}{\partial T_{B,i}} = a_i T_e^2 \frac{\partial \tau_i}{\partial T_{B,i}} \quad i = 1, 2 \quad (18)$$

$$= \frac{\pm k T_e^2}{2 w f_i^2} / (T_e - T_{B,i}), \quad \begin{array}{l} + \text{ for } i = 1 \\ - \text{ for } i = 2 \end{array}$$

Hence, the sensitivity is small for large $w f_i^2$ and small $T_{B,i}$ (well below saturation). For the typical meteorology profiles of Fig. 2, the sensitivities are calculated and summarized in Table 4. The frequency pair 22.235 GHz/31.4 GHz has the lowest sensitivity, as expected for its large w (cf. Table 1). The pair 18.0 GHz/21.0 GHz has a sensitivity a factor of 3 larger.

VII. Conclusions

The analysis indicates that it is possible to achieve a sub-centimeter estimate of water vapor path delay on nadir observation with an on-board microwave radiometer. To achieve such precision we need (1) a radiometer with two properly selected frequencies, (2) brightness temperature measurement accurate to ~ 0.5 K, (3) knowledge of sea surface temperature to ~ 3 K and (4) knowledge of sea surface wind speed to ~ 2 m/sec or knowledge of sea surface emissivity to ~ 0.01 . Note that the analysis does not apply when there is precipitation.

The frequency pair 24.3 GHz/32.0 GHz is found to be ideal for a nadir water vapor radiometer. It is insensitive to profile variation and has a moderate sensitivity to brightness temperature measurement error.

The frequency pair 18.0 GHz/21.0 GHz, though insensitive to sea surface temperature uncertainty, has high sensitivities to profile variation and to brightness temperature measurement error. It can be considered only when no knowledge of sea surface temperature is available and the brightness temperature measurement error is highly correlated (and thus cancelled) between the two frequencies.

References

- 1 Barrett, A H , and Chung, V K , "A Method for the Determination of High-Altitude Water-Vapor Abundance from Ground-Based Microwave Observations," *J Geophys Res* , Vol 67, No. 11, pp 4259-4266, 1962
- 2 Westwater, E R , "An Analysis of the Correction of Range Errors Due to Atmospheric Refraction by Microwave Radiometric Techniques," Tech Rep IER 30-ITSA 30, Environmental Science Services Administration, U S Dept of Commerce, Boulder, CO, 1967
- 3 Gaut, N E , "Study of Atmospheric Water Vapor by Means of Passive Microwave Techniques," Tech. Rep 467, MIT Res Lab of Electronics, Dec 20, 1968
- 4 Schaper, L W , Jr , Staelin, D H , and Waters, J. W , "The Estimation of Tropospheric Electrical Path Length by Microwave Radiometry," *Proc. IEEE*, Vol. 58, No 2, pp 272-273, 1970
- 5 Decker, M. T , Guiraud, F. O , and Westwater, E. R., "Correction of Electrical Path Length by Passive Microwave Radiometry," in *Proc. Conf on Propagation of Radio Waves at Frequencies above 10 GHz*, London, England, April 10-13, 1973.
- 6 Menuis, A C , Martin, C F., Layson, W M , and Flagg, R S , "Tropospheric Refraction Corrections Using a Microwave Radiometer," Tech. Staff Tech Memo, 19 ETV-TM-64-12, Pan American Airways, Nov 16, 1964
- 7 Wu, S C , "Optimum Frequencies of a Passive Microwave Radiometer for Tropospheric Path-Length Correction," *IEEE Trans Ant Prop* , Vol AP-27, No 2 , pp 233-239, March 1979
- 8 Bean, B R , and Dutton, E J , *Radio Meteorology*, National Bureau of Standards Monograph 92, Boulder, CO, 1966.
- 9 Grody, N C., Gruber, A , and Shen, W C., "Atmospheric Water Content Over the Tropical Pacific Derived from the Nimbus-6 Scanning Microwave Spectrometer," *J Appl Meteorology*, Vol 19, pp 986-996, August 1980
- 10 Staelin, D H , "Measurements and Interpretation of the Microwave Spectrum of the Terrestrial Atmosphere near 1-Centimeter Wavelength," *J Geophys Res* , Vol 71, No 12, pp 2875-2881, 1966
- 11 Van Vleck, J H., Purcell, E M , and Goldstein, H., "Atmospheric Attenuation", in *Propagation of Short Radio Waves*, D E. Kerr, ed , McGraw-Hill, Chapter 8, 1951
- 12 Wilheit, T. T , Jr , "A Review of Applications of Microwave Radiometry to Oceanography," *Boundary-Layer Meteorology*, Vol. 13, pp 277-293, 1978
- 13 Nordberg, W., Conaway, J , Ross, D B , and Wilheit, T , "Measurements of Microwave Emission from a Foam-Covered Wind Driven Sea," *J Atmos. Sci* , Vol. 28, pp. 429-435, 1971.
- 14 Webster, W. J , Jr , Wilheit, T T , Ross, D B , and Gloersen, P , "Spectral Characteristics of the Microwave Emission from a Wind-Driven Foam-Covered Sea," *J Geophys Res* , Vol. 81, No 18, pp. 3095-3099, June 1976.

Table 1. Values of T_{ox}^2 and $w(0)$ for typical meteorology profiles

f_1	20.2 GHz	24.3 GHz	20.3 GHz	22.235 GHz	18.0 GHz
f_2	26.0 GHz	32.0 GHz	31.4 GHz	31.4 GHz	21.0 GHz
T_{ox}^2	0.579	0.189	0.659	0.394	0.565
$10^3 w(0)$	1.03	1.11	1.60	1.89	0.99

Table 2. Calibration uncertainty of ΔR (cm) using theoretical coefficients derived from Table 1

f_1	20.2 GHz	24.3 GHz	20.3 GHz	22.235 GHz	18.0 GHz
f_2	26.0 GHz	32.0 GHz	31.4 GHz	31.4 GHz	21.0 GHz
Case 1	-0.06 ± 0.15	-0.24 ± 0.18	-0.26 ± 0.18	1.19 ± 0.64	1.72 ± 0.77
Case 2	-0.07 ± 1.44	-0.26 ± 0.98	-0.28 ± 1.25	1.19 ± 1.39	1.71 ± 0.80
Case 3	-0.06 ± 1.16	0.16 ± 1.13	0.28 ± 1.49	1.36 ± 1.06	1.39 ± 1.32
Case 4	-0.07 ± 1.94	0.13 ± 1.43	0.25 ± 1.85	1.37 ± 1.67	1.36 ± 1.43

Table 3. Calibration uncertainty of ΔR (cm) using regression coefficients derived from forced-fitting ΔR and the simulated τ through Eq. (10)

f_1	20.2 GHz	24.3 GHz	20.3 GHz	22.235 GHz	18.0 GHz
f_2	26.0 GHz	32.0 GHz	31.4 GHz	31.4 GHz	21.0 GHz
Case 1	0.00 ± 0.15	0.00 ± 0.15	0.00 ± 0.15	-0.02 ± 0.47	-0.02 ± 0.45
Case 2	0.00 ± 1.32	0.00 ± 0.96	0.00 ± 1.19	0.00 ± 1.13	0.00 ± 0.51
Case 3	0.00 ± 0.99	0.01 ± 1.07	-0.02 ± 1.21	-0.01 ± 0.71	0.01 ± 0.92
Case 4	0.00 ± 1.54	0.00 ± 1.31	0.00 ± 1.67	0.00 ± 1.22	0.00 ± 0.98

Table 4. ΔR sensitivity to T_B measurement error

f_1	20.2 GHz	24.3 GHz	20.3 GHz	22.235 GHz	18.0 GHz
f_2	26.0 GHz	32.0 GHz	31.4 GHz	31.4 GHz	21.0 GHz
$\frac{\partial(\Delta R)}{\partial T_{B,1}}$	1.17	0.79	0.75	0.57	-1.42
$\frac{\partial(\Delta R)}{\partial T_{B,2}}$	-0.73	-0.47	-0.34	-0.29	1.16
Worst case	1.90	1.26	1.09	0.86	2.58

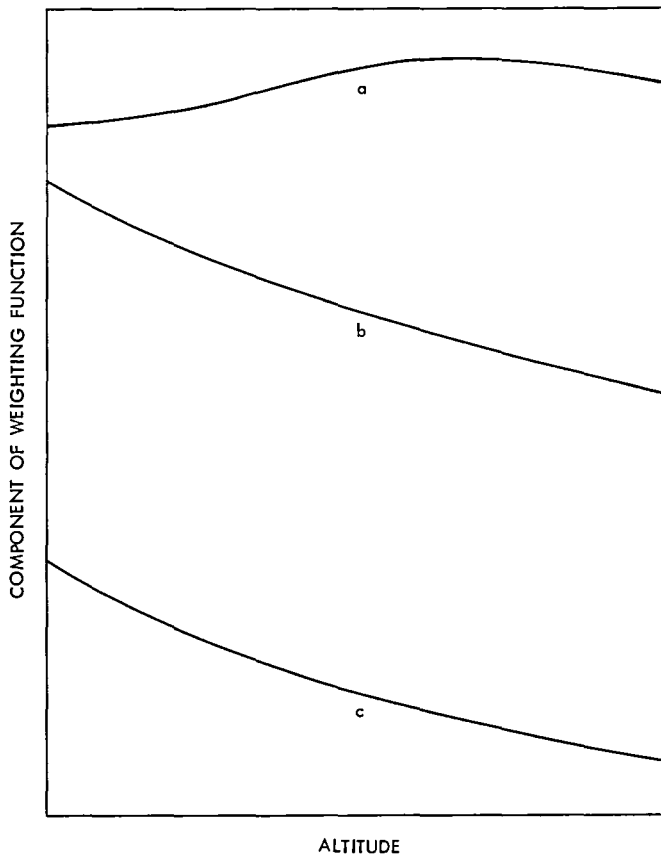


Fig. 1. Descriptive examples of weighting function components at three different frequencies; the pair combining b and c has a constant separation and will result in a constant weighting function

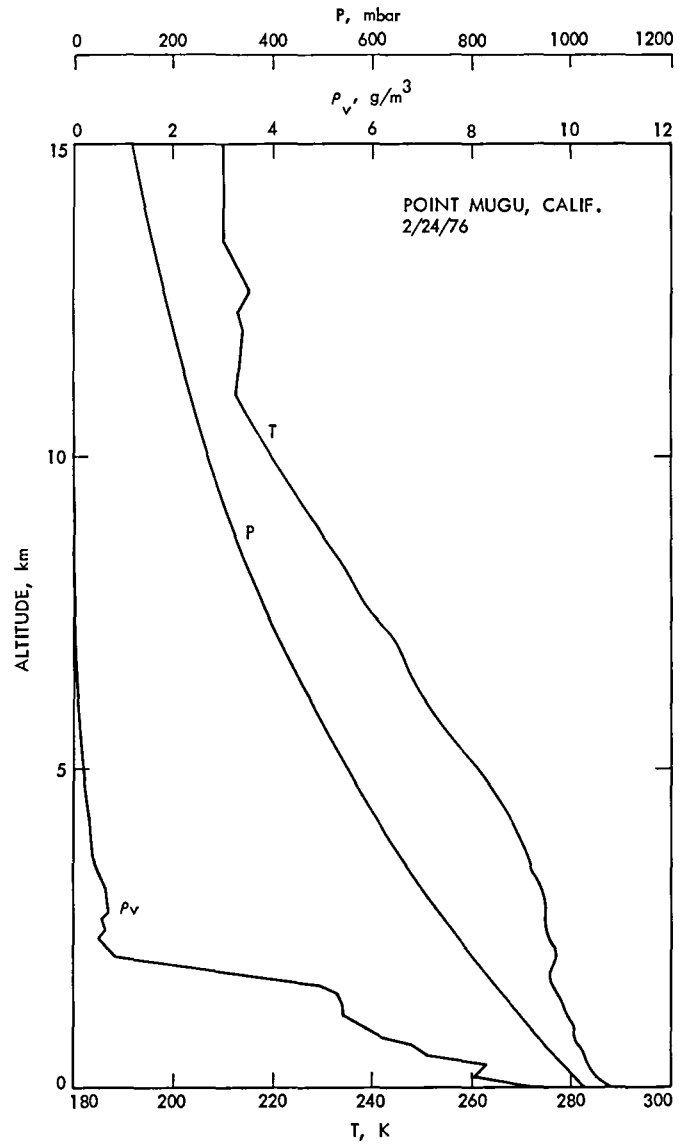


Fig 2. Typical profiles of atmospheric temperature, pressure and water-vapor density

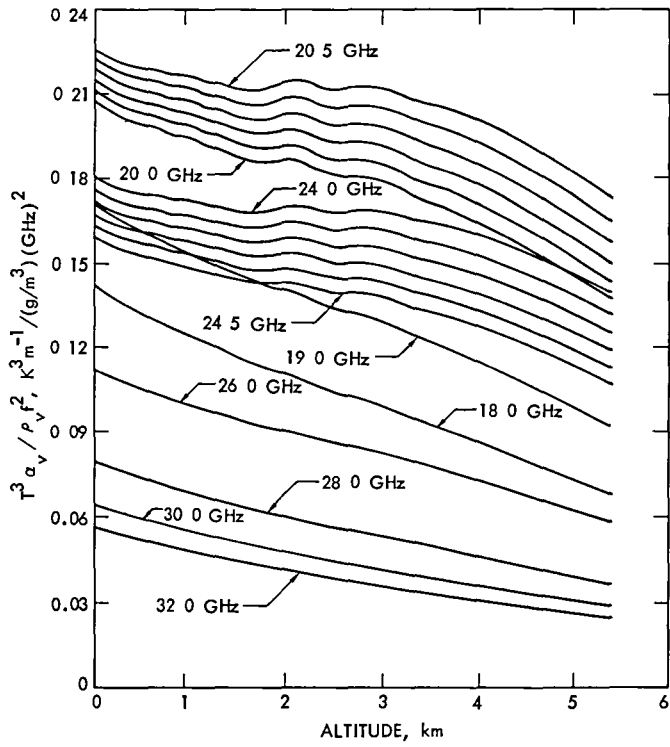


Fig 3. Components of weighting functions for various frequencies

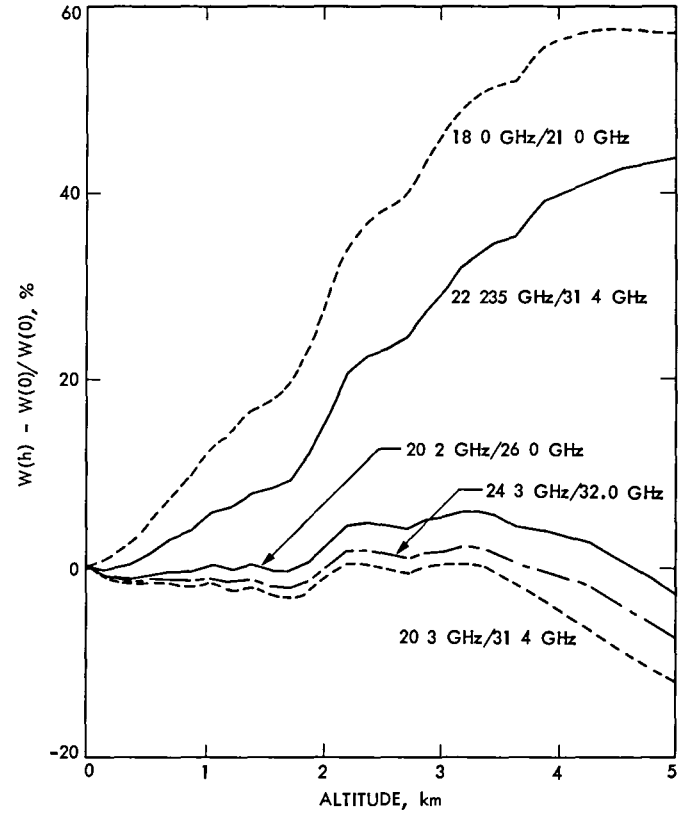


Fig 4 Normalized weighting functions for various frequency pairs

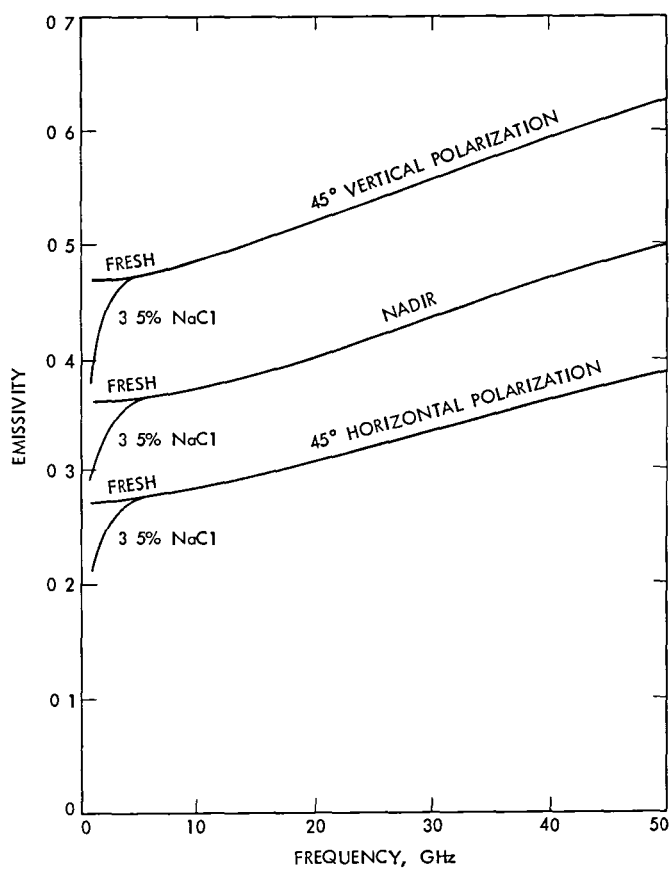


Fig 5. Emissivity of water at temperature $T_s = 290$ K (from Ref 12)

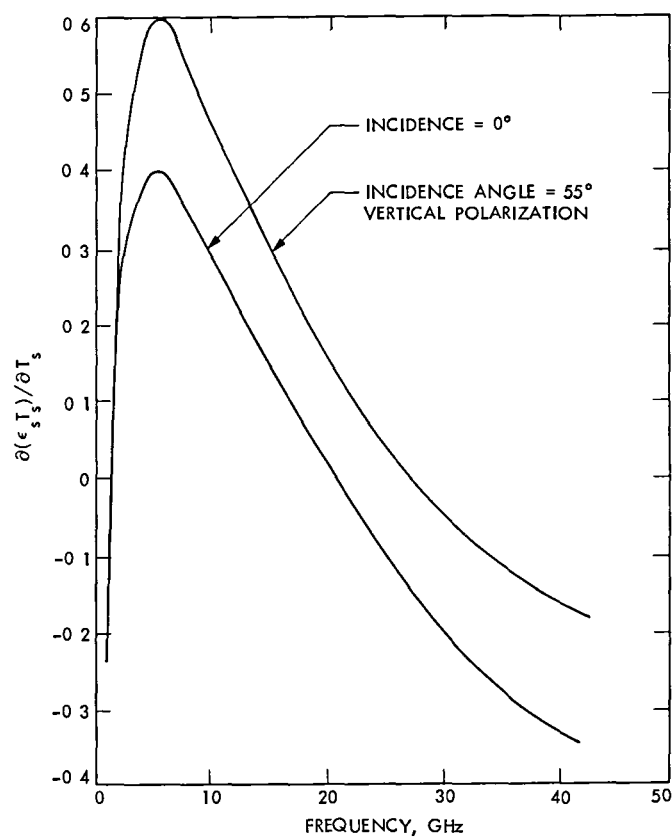


Fig 6 Change of sea surface emission ($\epsilon_s T_s$) per change of sea surface temperature (T_s) as a function of frequency (from Ref 12)

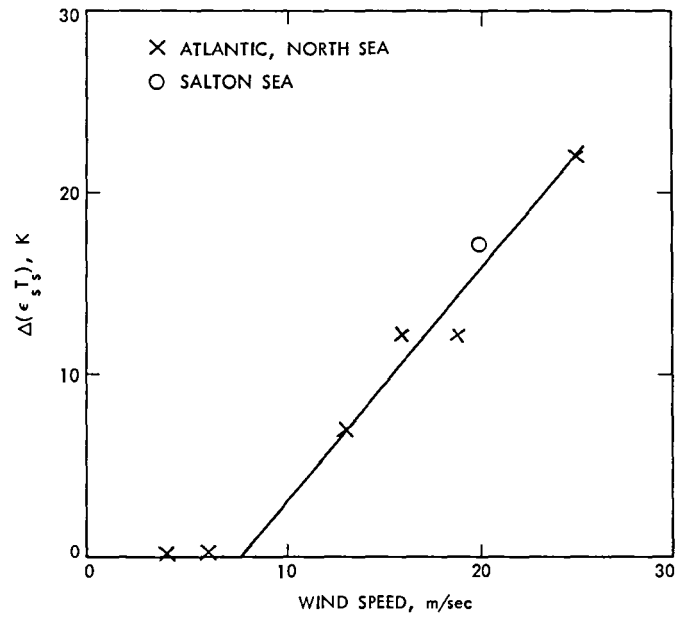


Fig. 7 Increase of $\epsilon_s T_s$ at 19.35 GHz caused by wind speed at sea surface (from Ref 13)

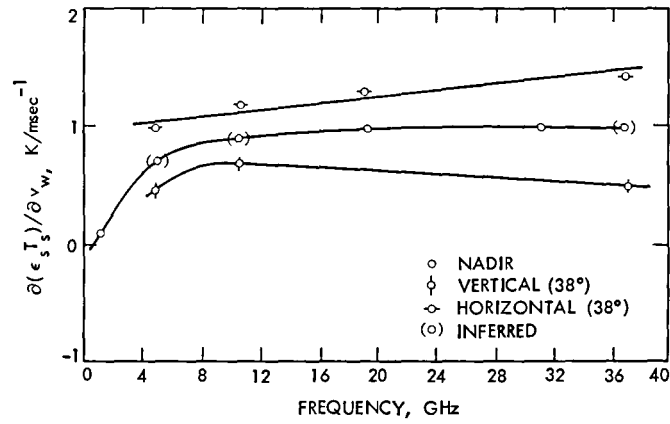


Fig. 8. Rate of increase of $\epsilon_s T_s$ with wind speed as a function of frequency (from Ref 14)

Block I Phase 2 Very Long Baseline Interferometry Implementation

J H Wilcher
DSN Data Systems Section

This article is a status report on the implementation of the Block I VLBI System

I. Introduction

The Block I VLBI System (Ref 1) is presently being implemented in the Deep Space Network as a phased implementation. The Block I Phase 1 VLBI System has been previously reported (Ref 2) Phase 2 of the Block I VLBI System implementation has been completed and is now operational

II. DSS Configuration

The DSS configuration for Block I Phase 2 is shown in Fig. 1 and remains unchanged from the Block I Phase 1 configuration. Planned changes in the near future to the DSS configuration include

- (1) The replacement of the digital magnetic tape Real-Time Record by a large capacity disc serving the same function This change obviates the requirement for the many tape changes during the VLBI passes The disk will be large enough to handle a complete pass
- (2) The inclusion of a digital tone extractor function to the DSS VLBI Subsystem. This addition will permit real-time verification of the presence of the phase calibration tones through the DSS VLBI System

III. Ground Communication Facility and Network Operations and Control Center Configuration

The GCF-NOCC configuration for Block I Phase 2 is shown in Fig 2. The Ground Communications Facility has undergone changes in configuration (Ref 3) The significance of the changes in GCF to the Block I VLBI System lies in the direct connection of the wideband data lines from the wideband data terminals to the recently installed Block I VLBI Processor Subsystem. This change allows the VLBI data to be routed directly from the DSSs to the VLBI Processor Assembly, where the data are recorded on large-capacity (600 Mbyte) discs The creation of numerous IDRs and the attendant handling of such is thus made unnecessary.

The Network Operations and Control Center configuration was changed by the addition of the Block I VLBI Processor Subsystem The incorporation of the Block I VLBI Processor Subsystem (VPS), which consists of the VLBI Processor Assembly, the VLBI Correlator Assembly and accompanying software, removes from the Block I VLBI System the Interim VLBI Processor Subsystem. The Interim VLBI Processor consisted of a correlation program on a large general-purpose computer (IBM 360-75). This configuration required the

generation of numerous IDRs in order to transfer the VLBI data to the processor along with the necessity of waiting sometimes many days for the processing to be completed. The addition of the dedicated Block I VLBI Processor Subsystem allowed the processing to be completed in a matter of hours after the receipt of the VLBI data from the participating DSSs.

IV. Correlation and Postcorrelation with the Addition of the VPS

The correlation of the VLBI data is now carried out on the VLBI processor with hardware and software well-documented and transferred to Operations. The postcorrelation effort is carried out on the same processor. However, the software involved is considered Research and Development in nature at the present time. An effort is underway to document and transfer the postcorrelation software to Operations at a later date.

V. Operation Training

DSS operations have been going quite well, with the VLBI passes being scheduled and carried out almost weekly for the past year and a half. The operational proficiency seems to indicate that the training packages were very adequate. The

training for the VLBI Processor Subsystem was carried out in a much more informal manner than the DSS training, since there were fewer personnel involved. The training consisted of theory of operation discussion and "hands on" operations with the Development and Operations personnel involved. The operation is currently completely carried out by the Operations personnel.

VI. System Testing

System testing was accomplished by operating the Block I VLBI Processor Subsystem and the interim processor in parallel. This was accomplished by creating IDRs of the same VLBI data that was being obtained by the VPS and then processing the data through both the VPS and the interim processor and verifying that the results were identical in every step of the processing. This test was repeated during several VLBI passes. In every case the processing was verified as identical.

VII. Summary

The Block I Phase 2 VLBI System is operating to support the navigation requirements of the Voyager Project. The System will provide the information related to station clock synchronization, UT1 and polar motion in a timely manner.

References

- 1 Chaney, W. D., and Ham, N. C., "DSN VLBI System MICI-80," in *The Deep Space Network Progress Report 42-56*, pp. 26-34, Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1980.
2. Wilcher, J. H., "Block I, Phase 1 Very Long Baseline Interferometry Implementation," in *The Deep Space Network Progress Report 42-58*, pp. 24-27, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1980.
- 3 McClure, J. P., "GCF-NOCC Reconfiguration," this issue.

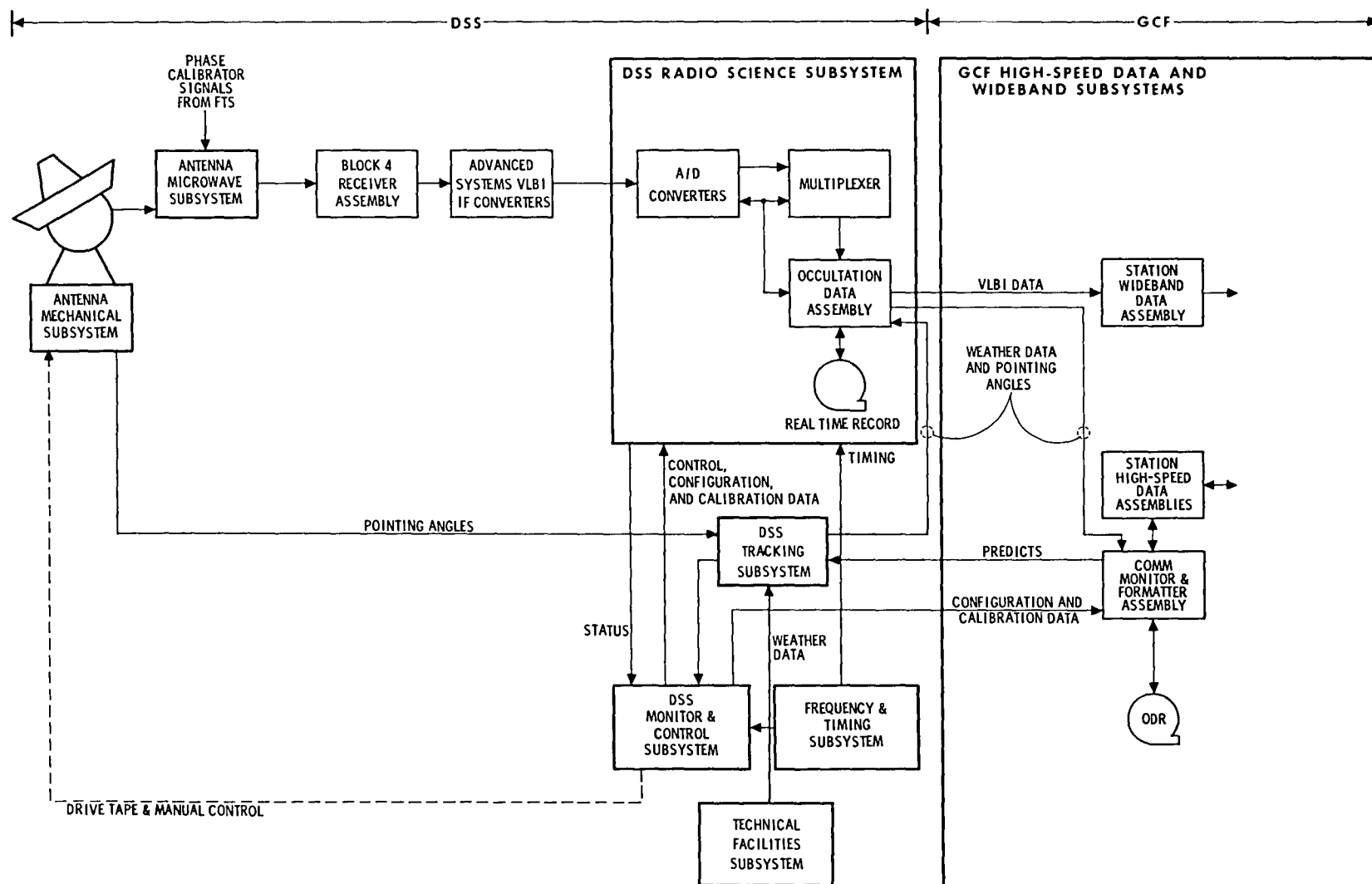


Fig 1. Deep Space Station block diagram, Block I Phase 2

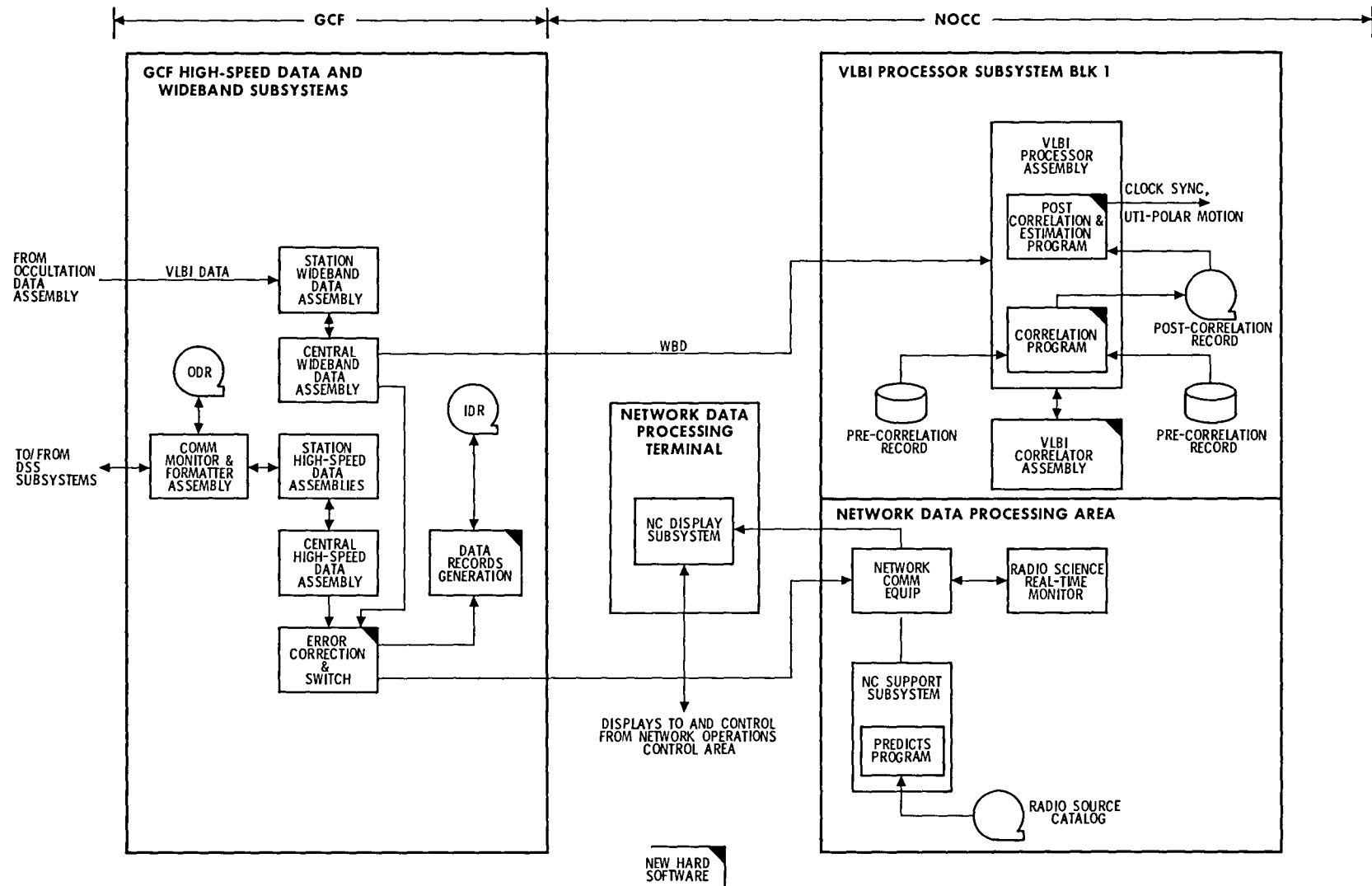


Fig. 2 GCF-NOCC block diagram, Block I Phase 2

The DSN Programming System

G. A. Madrid
TDA Engineering Office

This article describes the DSN Programming System and its current status. Accomplishments and plans are summarized. A study to revise the system design is in progress.

I. Introduction

The DSN Programming System is an interrelated body of procedures, tools, equipment, and information designed to assist DSN software developers and managers in the efficient and effective production of software for DSN tasks. The major goals of this system are (1) to reduce the cost of producing, documenting, and maintaining DSN software, and (2) to improve the ability of managers to cost, schedule, monitor, and evaluate DSN software-development tasks under their purview.

The DSN Programming System is of particular importance to the Mark IV-A Implementation Project due to its need to maintain strict budgetary and schedule constraints on all development being performed under their aegis. It is expected that the Programming System will provide just this type of assistance in the development of computer software for the new series of computers replacing those currently at the DSN deep space stations (MODCOMP II computers are being replaced by MODCOMP CLASSICS).

II. System Objectives

To achieve the stated goals, the DSN Programming System has set for itself the following objectives:

- (1) To provide implementing organizations the standard practices and software-development methodology

(Refs 1 and 2) necessary to increase the cost effectiveness of the software-development process. The increase will be reflected in reduced overall life-cycle costs due mainly to improvements in maintainability and reliability of software developed using common standards.

- (2) To develop and provide the standard languages and software-development tools needed to code, check out, and document computer programs for the DSN. These improvements in the developmental environment will result in improved productivity and personnel morale. These are major ingredients in developing software on schedule and within budget constraints.
- (3) To develop and provide computer-generated aids and reports needed to increase management visibility into the software-development process. This increased visibility will permit a more efficient use of available resources, a timely recognition of problems, and the information required to reduce the risk in software-management decisions.
- (4) To assure, by means of an appropriate functional design, accord of the Programming System with the actual software-development process. This will permit an evaluation of the system based on its functional flow and more clearly delineate the interfaces and functions required by the system to support the software-development process.

- (5) To develop the methodology and procedures needed to measure system performance against realistic performance criteria. This methodology needs to be a permanent fixture in any new system design if that system is expected to remain a viable entity for any length of time.

III. System Description

The elements and structure of the current DSN Programming System are shown in Fig 1. It is, in design and approach, very much the same system described by Hodgson (Ref 3) and reported by Irvine (Ref 4). The basic structure is defined in five basic categories:

- (1) Software Standard Practices
- (2) Standard Processors
- (3) Programming System Data Base
- (4) Management Aids.
- (5) Implementation Aids

This structure, however, does not provide a functional design for the system and therein lies its major deficiency. Section V discusses this problem and the study underway to redefine the system.

The Software Standard Practices segment of the system is represented by a series of documents that specify standard methodology, guidelines, and practices dealing with the implementation of software from inception to operations. These documents are maintained in the 810-XX series of DSN documentation. However, a compendium of these standards (Ref. 5) also exists and may be more suitable for those readers merely desiring to inform themselves regarding the DSN Standard Practices. The major activity in this area currently involves reviewing these standards in light of their effectiveness and usefulness to the implementing organizations. Modifications are made as required contingent on appropriate review board approval.

The Standard Processor segment of the system is responsible for developing and maintaining the standardized software environment needed by the software developers and managers of software-implementation efforts. This function was responsible for selecting and implementing the two current Standard Languages: MBASIC* and HAL/S. The former is termed the "off-line" language to be used for administrative and nonreal time applications. The latter is termed the "on-line" language to be used for real-time DSN applications. Procurement of

other useful tools and languages are also performed under the purview of this system element.

The Implementation Aids segment of the system is an adjunct to the Standard Processor segment. Its main purpose is to provide the software developers additional tools to assist them in designing, documenting, and testing computer programs. The Data Base and Management Aids segments are also complementary in that their purpose is to provide the management aids, reports, and information required to improve implementation process visibility and ensure informed management decisions at critical points in the development effort.

IV. Status and Plans

This section discusses the status and plans of those segments of the system where activity or changes occurred since the last report (Ref. 3).

A. Software Standard Practices

Standards for the production of System and Subsystem Functional Requirements are in draft form. Finalization of these documents is being placed in abeyance pending a formal review of the Programming System Functional requirements and design. Such a review is expected in the first half of calendar year 1982 (see Section V).

B. Standard Processors

The majority of activity in the Programming System has dealt with the implementation of the "on-line" language, HAL/S. This language has been identified by the DSN as the standard real-time language to be used in any software development for the real-time computers being installed at its remote stations under the Mark IV-A Implementation Plan.

A version of the HAL/S Compiler for listing on a MODCOMP 7870 computer was developed by Intermetrics, Inc., under contract to JPL. A completed version of this compiler was accepted by JPL on May 14, 1981. Delivery was accepted despite some errors found during testing of the product. Of the original errors found, only four were considered major problems, and all of these have been resolved by Intermetrics under a maintenance contract with JPL. There are currently 22 outstanding problems of which 2 are considered of major importance. All of these are due to be resolved by January 1982.

A MODCOMP 7870 with operating system and peripherals was purchased and installed in October 1980 to serve as the host development computer for the HAL/S language and its associated tools.

*A trademark of the California Institute of Technology.

Training on use of the new compiler language has begun. A training class conducted by Intermetrics was held on July 20, 1981, for 25 programmers and supervisors who will be utilizing the language in their implementation tasks. Additional classes are planned.

Development of a more comprehensive set of tests for the HAL/S compiler is underway. Results from these tests will be used as a condition for the transfer of the compiler to operational status. Such a transfer is not expected until the first quarter of calendar year 1982 when it is expected that the number of problems uncovered through testing will have stabilized. This date will not impact Mark IV-A software development since a code can be developed using the current compiler and then recompiled with the operational version when that is made available.

A list of HAL/S-compatible tools and aids available from other NASA centers has been compiled for use in determining how to improve the DSN development environment as quickly and as cost effectively as possible. Active participation of TDA and implementing organization personnel in the NASA HAL/S User and Configuration Control Board is being encouraged so that repetition of costly mistakes can be avoided through the sharing of information with HAL/S users at other NASA centers. This could lead to reduced development costs and reduce the impact of unnecessary language changes on the users.

The DSN Standard "off-line" language, used for nonreal time purposes, is being maintained for current users on the JPL institutional computers (Univac 1100/81). A study to replace this standard with ANSI-BASIC is being proposed. The decision for the computer host of this language should be subsidiary to the more fundamental decision of the functional design of the Programming System itself.

C. Management Aids

A preliminary version of a Work Breakdown Structure (WBS) report generation system has been made available to management. This management aid has elements on both the AODC (Automated Office Data Center, a micro-processor

computer system developed by JPL) and the MODCOMP 7870 currently being used for the HAL/S development. A Software Cost Estimation Model (SOFTCOST) is similarly hosted on the AODC with data access through the interconnection with the MODCOMP 7870.

Current planning is directed toward providing the management functions with a separate MODCOMP 7870 thus freeing the development computer for more on-line use by software developers. The management computer would then be the central node for the network of terminals and processors currently used by management to maintain data bases, estimate cost, report productivity, and produce up-to-date schedules and resource allocations.

V. Future Plans

The use of the Programming System to support the Mark IVA Implementation effort has indicated a need to review the design and structure of the current system.

The lower-level functional elements of the system are well defined and can be directly related to requirements or constraints. These lower-level elements, however, are not organized along functional lines but along generic lines. If the system is to support the Mark IV-A development effort, a better approach would be to organize these elements in conformance with the actual functional flow of the development process. This will permit a more conventional division of the system into subsystems as well as a clearer definition of the interfaces between subsystems and the interaction between elements. Such a redesign would permit a clearer evaluation of priorities and the impact of changes on the entire system would be more clearly discerned.

A study to propose a revised system design along these lines is underway and will be reviewed and presented in the first half of calendar year 1982. It is also planned to provide a long-term plan that will guide the system development for the next five years and provide some insight into future development through 1992.

References

1. Tausworth, R. C , *Standardized Development of Computer Software, Volume II Standards* Special Publication 43-29, Jet Propulsion Laboratory, Pasadena, Calif., Aug 1977
2. Irvine, A. P., and McKenzie, M., "Evaluation of the DSN Software Methodology," in *The Deep Space Network Progress Report 42-48*, pp 72-81, Jet Propulsion Laboratory, Pasadena, Calif., Dec 15, 1978.
3. Hodgson, W. D., "The DSN Programming System," in *The Deep Space Network Progress Report 42-41*, pp. 4-9, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1977.
4. Irvine, A P , "The DSN Programming System," in *The Deep Space Network Progress Report 42-50*, pp. 4-6, Jet Propulsion Laboratory, Pasadena, Calif , Jan – Feb 1979.
5. Irvine, A. P., ed., *Standard Practices for the Implementation of Computer Software*, Publication 78-53, Jet Propulsion Laboratory, Pasadena, Calif , Sept 1978

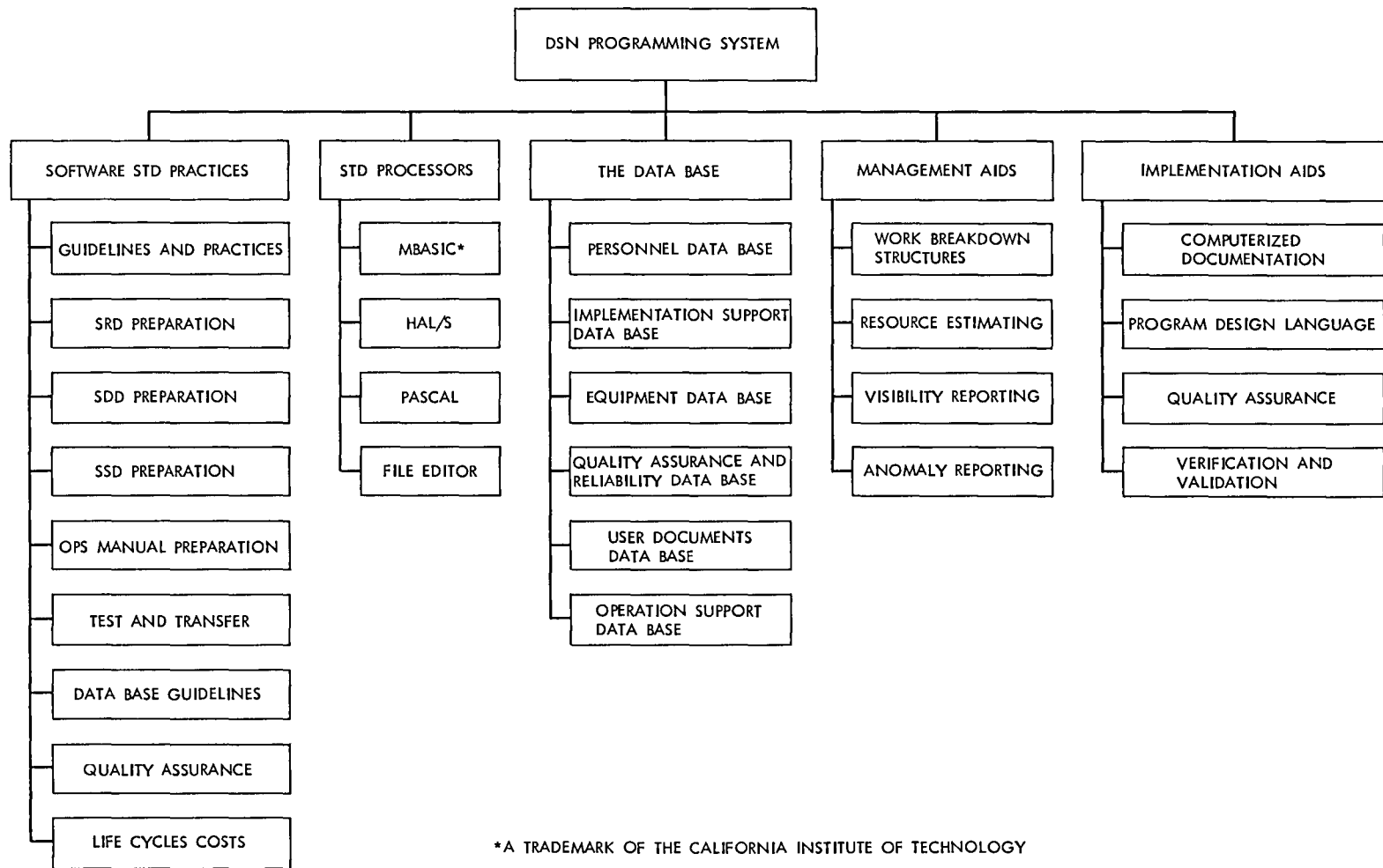


Fig 1. Structure and elements of the DSN Programming System

Voyager Mission Support (I)

N. Fanelli and H. Nance
Control Center Operations

This article is a continuation of the Deep Space Network report on tracking and data acquisition for Project Voyager. This article covers the period July through 11 August 1981.

I. Introduction

Voyager 1 was in the extended mission phase of operations during this reporting period. Voyager 2 completed the Observatory phase of the Saturn Encounter on 31 July and the Far Encounter-1 (FE-1) phase on 11 August. It is now in the Far Encounter-2 (FE-2) phase in preparation for the nearest approach on 26 August.

II. DSN Support

A. Voyager 1

The extended mission phase continued for Voyager 1, with the 26-meter-antenna net providing the majority of the tracking support. However, 34- and 64-meter support was provided as required, as on the Navigational Cycle (Nav Cycle) performed on 20/21 July. This Nav Cycle was performed with only three-station coverage; the previous cycles had employed four-station coverage (two tracks by the Australian stations). The tracking loop capacitor (TLC) test continued to be run weekly as was the transmission of a dummy computer command (CC) to prevent the spacecraft from going into the emergency routine when a command had not been received for a specified period.

Radio Science return was received during the XSCEL (X-band Saturn Celestial Mechanics) activity in the form of

Saturn system celestial mechanics study, gravitational red shift determination, and ultrastable oscillator frequency check. The activity was conducted biweekly. Periodic Engineering and Science Calibration (PESCAL) activities were supported bimonthly.

B. Voyager 2

1. Operational Readiness Test B-2 On 1 July, the Radio Science Operational Readiness Test B-2 was conducted with DSS 43. A command load was put into the spacecraft to simulate the closest approach conditions so that the encounter Radio Science sequence could be exercised to validate procedures and end-to-end support system operation. The test was very successful and the major objectives were met. Some minor problems were identified that required corrective actions. On 5 August, a mini-ORT was conducted to test the system again and verified the corrections and procedures as applicable for the problems identified in the ORT.

2. Radio Science Operational Verification Tests During the reporting period, DSS 43 Radio Science capabilities were exercised by conducting five Operational Verification Tests (OVTs). The OVTs primary objectives were to insure that the equipment remained operational, configurations were valid, and personnel were trained and ready to support the Radio Science activity during the Saturn closest approach. Although some minor problems were experienced during the

testing, indications generally were that the equipment and personnel were ready to support the actual encounter operation satisfactorily

The primary problem was the implementation of the Precision Power Monitor (PPM) capability to monitor the system noise temperature (SNT) of all four receiver systems (two S-band and two X-band). Although the system was installed, continued proper operation could not be achieved. Two engineers from JPL were dispatched to the station to assist in correcting the problems.

DSS 43 also experienced problems with the Radio Science wideband recorders due to excessive bit error rate (BER). The problem was isolated and corrected by the use of high quality tapes. With proper alignment, new analog-to-digital converters were installed and the operation validated during operational testing.

It was felt that some of the recorder problems could be attributed to the recording tape being used. To overcome this possibility, special, validated and certified tapes (Ampex 799) were shipped to the station to be used for specified test periods and for the actual Near Encounter Radio Science activity. Use of the tapes has generally improved recorder operation.

3 Trajectory Correction Maneuver B-8 On 19 July, a trajectory correction maneuver was performed by Voyager 2 to position the spacecraft at Saturn closest approach at approximately 161,000 km from the center of the planet. The maneuver was executed by roll and yaw turns to place the spacecraft on the burn axis, which placed the spacecraft off earth-point. The spacecraft was off earth-point for two hours and 38 minutes, during which no downlink was obtained.

DSS 12/42 supported the preburn activity, which included special commands and maneuver enable commands. The roll/yaw turns were executed over DSS 42 and the downlink was lost on time. The actual burn occurred during DSS 63/61 view period. The burn was accomplished and the spacecraft returned to earth-point as programmed. To assist in the reacquisition of the X-band downlink, DSS 63 had initialized its Occultation Data Assembly (ODA) and Spectral Signal Indicator (SSI) to search the spectrum for the first indication of the signal. The signal was observed on the SSI and by the X-band receiver, and DSS 63/61 reacquired the downlink telemetry on time.

Critical commanding followed the reacquisition of the spacecraft downlink, for included in the onboard computer was a Maneuver Recovery Block (MRB) to be executed in case of an anomaly and the spacecraft did not return to earth-

point. The MRB would cause the spacecraft to perform a 360-degree roll so that the signal strength could be plotted and the spacecraft pointing could be determined. Once the final pointing position was determined, the spacecraft could be commanded to return to earth-point. To support this activity, DSS 63 had counted down the 100-kW transmitter to be used in case the MRB was executed and the extra uplink power was required to command the spacecraft through the low-gain antenna. The critical commanding was to disable the MRB maneuver within a limited command window. To accomplish the commanding, DSS 63 was required to establish an uplink ramp that would cross through the bandpass of the spacecraft receiver. The spacecraft has a tracking loop capacitor failure that reduces the receiver bandpass, and the receiver frequency varies with changes in compartment temperature. While DSS 63 was ramping the uplink, 30 MRB disable commands were transmitted on 1-minute centers. DSS 61 was ramping also, with the transmitter off, as a backup to DSS 63. The DSS 63 command system alarmed on the first command, and DSS 61 immediately turned its transmitter and command system on, and continued commanding five minutes later. The disable commands were received by the spacecraft and the MRB was disabled. This emergency procedure is documented and is used in support of each spacecraft maneuver.

4 Navigation Cycles Dual Nav Cycles were supported by the 34/64-meter nets preceding and following the TCM B-8 (see above) to provide the basis and verification of the results of the burn. During a dual Nav Cycle, seven station tracks are provided, whereas during a standard Nav Cycle only four station tracks are provided. The Nav Cycles begin in Australia (southern hemisphere) and continue with Spain/US (northern hemisphere). The Australian stations support the first, middle and last tracks, providing two southern and two northern tracks per cycle.

Three standard Nav Cycles were supported during the period to provide Project Navigation updated data for orbit determination. During processing of the Nav Cycle range data it was discovered that DSS 43's data had a 1-second timing bias. It was necessary for Project Navigation to compensate for the error during processing. The range data from DSS 14 and 63, as well as the doppler data from all three stations, taken during the cycle were good and met Project requirements.

5 Observatory Phase The Observatory Phase ended at 1003Z on 31 July and the FE-1 phase started. The imaging data received in real-time and during replay were of excellent quality and only a few problems were encountered. DSS 14 had sole plate problems with the antenna, causing the loss of 11 minutes of ultraviolet spectrometer mosaic data of Saturn. DSS 12 experienced a hail storm that ruptured the mylar shield on the S-band horn, but resulted in no impact to the

project. Bad weather in Australia caused the loss of four images severely degraded, and 27 images with moderate degradation at DSS 43.

6 Horizontal and vertical scan maneuvers These maneuvers were conducted from 31 July to 1 August. The purpose of the maneuvers was to align the ultraviolet scan field in a vertical position in respect to Saturn, allowing maximum resolution in the horizontal direction as defined by Titan's orbit and then in the horizontal position for the same purpose. The scan platform made a series of slews in each position to scan across the Saturnian system to study the atmospheric emissions.

The horizontal maneuver was supported by DSS 63/61 and the spacecraft stayed on earth-point throughout the maneuver so that telemetry data was received by the stations throughout the period. During the vertical maneuver the spacecraft was off earth-point for six and a half hours. The period up to the loss of the downlink was supported by DSS 14/12, and the reacquisition of the telemetry downlink was

supported by DSS 43/42. Included in the DSS 43/42 track was the MRB NO-OP (nonoperational) activity. This activity was the same as that performed after the TCM B-8 (see above), but was accomplished in this case by DSS 43/42. The activity was successfully performed by DSS 43 and the backup capability of DSS 42 was not required.

III. DSN Capabilities

Most of the remaining hardware and software implementation planned for the support of the Saturn Encounter was completed during the first week of July. Also completed were the DSN Operational Verification Tests, proficiency OVTs will continue through the first part of August. The major improvements included installation of X-band low-noise masers and rework of the antennas at the 34- and 64-meter stations. The notable software improvements were for the Metric Data Assembly (MDA) and Occultation Data Assembly (ODA) to support the uplink operation and radio science Near Encounter experiment.

Voyager Mission Support (II)

N Fanelli and H Nance
Control Center Operations

This article is a continuation of the Deep Space Network report on Tracking and Data Acquisition for Project Voyager. This article covers the period of 12 August through September 1981.

I. Introduction

Voyager I was in the extended mission phase of operations during this reporting period. Voyager 2 completed the Far Encounter-2 (FE-2), Near Encounter (NE), and Post Encounter 1 and 2 (PE-1 and -2) phases of the encounter operations and entered the Uranus Interstellar cruise phase.

II. DSN Support

A. Voyager 1

During the period covered by this report, the support provided Voyager 1 activities was accomplished primarily during the viewperiod of Spain and by DSS 62. During the Saturn closest approach celestial mechanics data were obtained by DSS 11, 42, and 44 to assist in the calibration of the Voyager 2 radio science data. One navigation cycle (Nav Cycle) was completed by the 64-meter net during the period of 26-27 September.

B. Voyager 2

1. **Ring movie** On 13 August, a series of ring images were started during the pass of DSS 61/63 and continued around the network for several days. The images were used to produce the ring movie under better lighting and approach angle than

that obtained during the Voyager 1 approach. The sequence was scheduled for some 31 hours and 42 minutes, including a playback period. The 34/64 meter stations were arrayed for this activity and the real-time data were received at the 44.8-kbps rate, with the required playbacks at a rate of 29.8 kbps. All through the Far Encounter period, high-rate imaging data were very successfully obtained nearly continuously by the DSN. This imaging included Saturn, its rings and the many Saturnian satellites. Dual DODRs were produced by the tracking 64-meter station to insure maximum data return. The sequence was completed and the images were considered of excellent quality.

2. **TCM B-9.** The final pre-Saturn Encounter trajectory correction maneuver (TCM B-9) was successfully supported on 18 August 1981. This maneuver was required to make the final correction so the spacecraft would arrive at the Saturn Encounter target point on-time per the mission requirement. A 1.018-m/sec delta-V change was required to achieve this, and was accomplished by a 380-sec motor burn. The actual burn was accomplished over DSS 12/14. The spacecraft performed a roll and yaw turn to place it in the correct position for the burn trajectory. This placed the spacecraft antenna off earth point and the downlink was lost for 1 hour and 17 minutes. DSS 14 reacquired the downlink at the predicted time and the data evaluation indicated that the burn was successful. Thus

was also confirmed by comparison of the orbit determination data taken during the navigation cycles conducted prior to and following the TCM

3 ORT-B3 On 19/20 August, the final Radio Science Operational Readiness Test (ORT-B3) was conducted with DSS 43. All the encounter required equipment had been installed and was operational and the finalized operational plan was used for the ORT. The test was successfully supported by the station with only a few very minor problems. The DSN was declared ready for the Saturn Encounter.

4 Saturn Near Encounter. The Near Encounter mission phase started on 25 August 1981. The recording of Celestial Mechanics data started on 15 August and extended through the Near Encounter to 10 days afterwards. Gravity fields of the planet, rings and satellites were mapped as well as data on gravitational redshift and ultrastable oscillator frequent stability. The 64/34-meter nets supported this activity in the arrayed configuration with the data being successfully recorded at the 64-meter station utilizing closed-loop doppler and range tracking.

The Radio Science Near Encounter Saturn occultation data and ring scattering data were recorded by DSS 43 utilizing both the open-loop (medium-band) and closed-loop systems. The data consisted of Saturn atmosphere and ionosphere information and microwave scattering properties of the rings at oblique angles.

High-rate imaging data of the closest approach images of Saturn and several of its satellites were very successfully obtained by the DSN. Most of these images were obtained with the narrow angle camera at 44.8 kb/s.

Concurrently with these high activity science operations, the spacecraft was maneuvered so that the directional properties of plasma fields and particles in near-Saturn environment could be characterized. The maneuvers were also structured so that the spacecraft accomplished such other goals as establishing Miaplacidus as a roll reference star, aligning the scan platform to obtain dark limb/bright limb observation, F-ring imagery, and limb tracking, acquiring Vega as the roll reference star, and aligning the spacecraft yaw axis with the Saturnian rotational axis for information on plasma flow and particle fluxes in directions other than those normally viewed. These activities were sequenced throughout the closest approach phase and were successfully executed (except for the last maneuver) and supported by the 64/34-meter nets. The last maneuver was cancelled due to a spacecraft problem.

After the near-encounter activities were executed and the spacecraft exited from Saturn occultation, it was discovered

that the scan platform was stationary in a position away from the desired pointing at Saturn and the rings. Only black sky image frames were being received. The Project went into a troubleshooting and investigative mode and the DSN geared for schedule and SOE changes in support of the Project activities. Primarily the last sequenced maneuver was cancelled and the first post encounter images were not received.

Initially, the scan platform position was determined and limited changes in elevation and azimuth were attempted. It was determined that the changes in elevation were normal, but the changes in azimuth were erratic and slow in one direction and less so in the other direction. Playback of data indicated that the scan platform had functioned properly while the spacecraft was occulted, and that it faulted just prior to egress. The platform was repositioned through a short series of stepped moves so that the camera was pointed back at Saturn's rings as the spacecraft continued away from the planet. Ring images were received in this mode, although it was not the planned sequence; the data were deemed satisfactory. The DSN supported this entire nonstandard period with no significant problems.

5 Phoebe movie One of the primary Post Encounter mission objectives was to obtain pictures of the satellite Phoebe. Rather than make extensive movements by the scan platform alone, it was decided to maneuver the spacecraft to star reference Canopus and make minor step movements of the scan platform to center Phoebe in the field of view. The Phoebe movie sequence was completed successfully in this mode, with the supporting DSN stations providing excellent data. At the completion of the image sequence, the spacecraft was maneuvered again to star reference Miaplacidus and the scan platform stepped to a safe position. During the spacecraft maneuver, the supporting stations were required to configure for a ramp, commanding operation for the Maneuver Recovery Block No-Op sequence. This activity was successfully executed by the DSN.

6 Saturn medium-band occultation data DSS 43 was required to make duplicates of the Medium-Band Digital Original Data Records (MB DODR). These duplicates were used to produce Intermediate Data Records (IDR) for Radio Science processing. The original DODRs were held at the station for further use as required in accordance with the Radio Science Operations Plan.

Data reduction from the IDRs was performed at CTA-21. The ODR validation consisted of two steps, the first was a dump of the header information for each record. This gave information as to station, spacecraft, time, record count and configuration. It also gave an indication of data loss due to weak spots in the ODRs or malfunctions in the record-reduc-

tion process. The second step is to perform fast Fourier transforms (FFT) and do a power spectrum analysis for all data on each tape. This gave the approximate signal frequency, signal-to-noise ratio and the character of the recorded passband. Comparison of these values with expected levels gave an indication of the value of the data for scientific analysis.

Preliminary evaluation of the first processed duplicate IDRs indicated that the original ODRs should be sent from the station to CTA-21 for reduction. In general, the quality of the data was reasonable, however, there was concern about the consistent pattern of data dropouts, with one tape in particular showing an unacceptable number of dropouts. Passband frequencies and SNRs were as expected and the passband was clean. Under strong signal conditions, DRA recorder B appeared to be of slightly better quality than recorder A. The original DRA recordings are expected to be of better quality than the duplicates, as indicated by past experience, and therefore will be used for all final data reduction.

7. Near Encounter Operation. The DSN support of the Near Encounter Operation was accomplished without any significant fault. In addition to the normal science DODRs provided by the 64-meter net, a backup DODR was provided by both the 64- and 34-meter stations. During the closest approach, DSS 43 generated 10 medium-band and 40 wide-band radio science DODRs, which will be used to produce 484 radio science IDRs. Image reception in the arrayed configuration was of excellent quality and no images were lost, due to DSN operation. Playback of the image data taken while the spacecraft was occulted was accomplished as scheduled and the image data before the scan platform malfunction were likewise of high quality. Even with the postoccultation platform problem, the mission objectives were essentially met and DSN support considered excellent.

8. TCM B-10. Trajectory correction maneuver B-10 was accomplished by the spacecraft on 29 September 1981. The maneuver occurred during the viewperiod of Goldstone and was supported by DSS 14, with DSS 12 as backup. DSS 12 was diverted from support of Voyager 1 as backup support to DSS 14 during this period due to a low film height problem experienced at DSS 14, which caused the antenna to go to

“brake” due to an alarm on Pad 3 (left front). DSS 14 began support again in a few minutes. The maneuver was designed to place the spacecraft on the proper trajectory for the 1986 Uranus Encounter. The spacecraft was turned to the proper orientation for the correction vector and a “burn” of 5754 seconds accomplished. This sequence placed the spacecraft in an off earth-point for approximately 2 hours, 40 minutes during which no downlink was available. The return to earth point occurred on time as a result of the programmed unwind and the downlink reacquired. Vega was acquired after the maneuver for star reference. DSS 14 ramp-commanded a low gain antenna (LGA) select No-Op, which prevented the spacecraft from going into the contingency mode that would have occurred if the maneuver had not been successful. The maneuver was successful and the primary evaluation indicates the “burn” was nominal. The next TCM will not take place until January 1985.

9. Voyager Uranus Interstellar Mission (VUIM). On 28 September 1981 the Post Encounter phase of the Saturn Encounter was completed and the VUIM phase started. Spacecraft activities affecting the scan platform will be severely curtailed for several months as analysis continues on the problem that affects platform motion in azimuth (side to side). The platform will be operated only at low rates of speed and over a limited range from 180 to 270 deg azimuth. This range gives a satisfactory set of positions for Post Saturn, Uranus and Neptune observations. Special configurations or operational procedures may be required to support future scan platform activities.

C. DSN Capabilities

On 9 September 1981, the DSS 14 hydrostatic bearing runner joints failed, and DSS 14 was unable to move the antenna to point for support of the Voyager 2 pass. This required rescheduling of DSN support while DSS 14 was down for hydrostatic bearing runner repairs. DSS 12 and 13 were primarily used to support the Voyager requirements, causing a minimum of impact to the Voyager and other projects.

The hydrostatic bearing runner joints were repaired and DSS 14 was able to resume normal tracking operations on 25 September 1981.

A VLBI Survey of the Northern Polar Cap Region

D. D. Morabito, R. A. Preston, and J. Faulkner
Tracking Systems and Applications Section

A VLBI survey at 2.29 GHz of the northern polar cap region (declination $> 69^\circ$) has been conducted using a baseline consisting of the NASA Deep Space Network (DSN) sites at Goldstone, California and Madrid, Spain. The purpose of this survey was to identify sources in the northern polar cap region which possess milliarcsecond components. High declination VLBI sources provide valuable geometric coverage for determining the spin axis components of baselines. Out of 48 candidate sources selected from the Bonn 5-GHz survey, 42 were detected to have compact structure on the California/Spain baseline at the 2.29-GHz observing frequency. This polar survey is part of a more general VLBI sky survey.

I. Introduction

Up to the present time, there has been an insufficient number of known high declination celestial radio sources that have strong milliarcsecond components suitable for VLBI observations. This paper presents the results of a recent search of the northern polar cap region for such VLBI sources. These results are part of a general sky survey for VLBI sources sponsored by the Deep Space Network Advanced Systems Program (Ref. 1).

High declination VLBI sources provide valuable geometric coverage for determining more accurate estimates of the spin axis components of baselines in VLBI programs aimed at clock synchronization, earth rotational orientation, and geodetic baseline measurements. The deficiency of high declination VLBI sources was due to the lack of high-frequency single antenna surveys in the northern polar cap region (declination $> 69^\circ$). Such surveys permit sources to be identified that might possess compact structure at the milliarcsecond level.

A high-frequency single-antenna survey involving several sources from the 69- to 90-deg declination zone was recently

performed at the 100-m antenna at Bonn, Germany, at 4.9 GHz (Ref. 2). Candidate high declination VLBI sources were chosen from this survey based on the following criteria:

- (1) Total flux density at 4.9 GHz was greater than 0.5 Jansky
- (2) Spectral index from 2.7 to 4.9 GHz was greater than -0.6

There were a total of 48 candidate sources selected from the Bonn survey. Of these sources, 17 had total flux densities at 4.9 GHz greater than 1 Jansky, and 31 had total flux densities between 0.5 and 1.0 Jansky. These 48 sources were then observed during three different observing sessions in March 1980.

II. Experiment Configuration

The VLBI observations were performed on March 3, 12/13, and 28, 1980, using a baseline consisting of a 26-m telescope (DSS 13) at Goldstone, California, and a 64-m telescope (DSS 63) at Madrid, Spain. The observations were performed at 2.29 GHz with right circular polarization. The receiver chain

consisted of an S-band traveling wave maser followed by a phase-stable S-band receiver (VLBI receiver at DSS 13 and BLOCK IV receiver at DSS 63) which converted the signal to an IF of 50 MHz. The Mark II (or BLOCK 0) VLBI recording system then recorded a 1.8-MHz sideband by digitally sampling at a 4-Mb/s rate. Phase stability of the receiver chain and digital sampling were controlled by hydrogen maser atomic clocks at both stations. Each observation of a source was 3 minutes in duration, with most sources being observed twice. The Goldstone-Madrid interferometer has a length of 8396 km corresponding to 64 million wavelengths at 2.29 GHz. This corresponds to a fringe spacing of 3 milliarcseconds.

III. Data Reduction

Matching tapes from both stations were then cross-correlated using the Caltech/JPL Mark II VLBI processor. Computer manipulation of the correlator output yielded the correlation coefficient for each observation or the fraction of bits on the two tapes that were correlated. The correlation coefficients were then converted into correlated flux density or the VLBI strength for each celestial radio source observation, using the procedure discussed in a previous paper (Ref. 1), with a scaling constant of 2.5 ± 0.1 .

The 5-sigma detection limit for each observation was about 0.1 Jansky. Correspondingly, the uncertainty in detected source strength due to random noise was about 0.02 Jansky. However, in practice, systematic errors at about the 8% level dominate the random contribution for most sources.

The tapes must be correlated over a range of relative tape delay and delay rate offsets to compensate for a priori source position uncertainties. Appropriate searches in these parameters were performed so that the sky was completely searched within 30 arcseconds of all nominal source positions. The 6 sources not detected are assumed to have been within the range of position searching.

The measured delay and delay rate offsets for detected sources allowed the source positions to be determined to about the 0.5-arcsecond level. Instrumental biases were removed by observing sources with accurately known positions throughout each experiment. Atmospheric contributions to delay and delay rate were modeled by monthly mean values (Ref. 3) and earth rotational position was determined according to standard Bureau International de l'Heure (BIH) values of UT1-UTC and polar motion. Ionospheric effects were not modeled and are assumed to be the limiting error source. An uncertainty of 0.4 arcseconds was quadratically added to all error estimates to account for unmodeled effects. The tech-

nique of reducing the delay and delay rate observables into source positions is discussed in another paper (Ref. 4).

In addition to measurements of correlated flux density and source position, measurements of total flux density were performed at DSS 13 for most sources using a noise adding radiometer (Ref. 5). The uncertainty in the total flux density measurements is mostly due to the 0.03 kelvin errors in the measurement of system temperature, but in the case of the strong sources is dominated by the 3% errors in the value of antenna sensitivity (K/Jansky). The 0.03 K temperature uncertainty translates to a 0.3 Jansky uncertainty in total flux density. Multiple averaging of total flux density estimates results in decreased uncertainties when multiple observations are available.

Antenna sensitivity at each site was determined by using observations of flux calibration sources (Ref. 6). Pointing errors were kept to a negligible level by boresighting a few strong sources in the northern polar cap region.

IV. Results and Discussion

Of the 48 northern polar cap sources on which VLBI observations were performed, 42 were found to have detectable milliarcsecond structure. The detailed results appear in tabular form in Table 1. Notes regarding table entries appear in Table 2.

A correlated flux density histogram appears in Fig. 1. A total of three sources had correlated flux densities of greater than one Jansky, while a total of 20 sources had correlated flux densities of greater than 0.4 Jansky. A sky distribution plot of the detected 42 sources appears in Fig. 2. Here the average value of the correlated flux density was used to determine the source strength symbol used in the plot.

Since the declinations of these sources are all above 69 degrees, they are visible 24 hours a day from the antennas used in these observations. This allows a wide range of spatial frequencies to be sampled. Detectable variation in the measured correlated flux density as a function of observation time might be observable due to source structure effects. Therefore, all but six candidate sources were observed at least twice, with the observations occurring at widely different hour angles whenever possible. The following sources exhibited significant three-sigma variation in correlated flux density between observations, indicating complex structure: 0604+72, 0740+82, 4C 71 07, 0950+74, 1039+81, 1044+71, 1058+72, 3C 309.1, 1637+82, 1749+70, 1803+78, 1928+73, 1946+70, and 2007+77. Source structure may be an important consideration for high accuracy VLBI measurements, and the effects

of sources such as these on interferometric phase should be understood when selecting candidate VLBI sources

Ten of these 48 northern polar cap sources were recently observed by Waltman et al (Ref 7) on a baseline having comparable resolution, but at a higher observing frequency (5 GHz)

V. Conclusion

Forty-two sources with declinations greater than 69 degrees were found to have compact structure at 2.29 GHz on the Goldstone/Madrid VLBI baseline. These northern polar sources will be useful for VLBI experiments in which determination of the spin axis components of baselines is important.

Acknowledgments

We are indebted to the instrumentation and observing assistance supplied by L J Skjerve of JPL, B Johnson of Ford Aerospace and Communications Corp, F Alcazar, B Luaces, and D Munoz of the Madrid Deep Space Communications Complex, and the Radio Science Support group at the Goldstone Deep Space Communications Complex. We would also like to thank the station directors at the DSN stations for supporting our program, namely, E Jackson (DSS 13) and J Fernandez (DSS 63)

References

- 1 Preston, R A , et al , "Establishing a Celestial VLBI Reference Frame-I Searching for VLBI Sources," *The Deep Space Network Progress Report 42-46*, May and June 1978, Jet Propulsion Laboratory, Pasadena, Calif , Aug 15, 1978
- 2 Kuhr, H , Nauber, V , Pauliny-Toth, I I R , Witzel, A , "A Catalogue of Radio Sources", Max-Planck-Institut fur Radioastronomie, Preprint No 55, August 1979
- 3 Thuleen, K L , and Ondrasik, V J , "The Repetition of Seasonal Variations in the Tropospheric Zenith Range Effect," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol VI, pp 83-98, Jet Propulsion Laboratory, Pasadena, Calif , Dec 15, 1971
- 4 Morabito, D D , et al , "Arcsecond Positions for Milliarcsecond VLBI Nuclei of Extragalactic Radio Sources, Part I 546 Sources," to be published in the *Astronomical Journal*, 1981
- 5 Reid, M S , Gardner, R A , and Stelzried, C T , "A New Broadband Square Law Detector," JPL Technical Report 32-1599, Jet Propulsion Laboratory, Pasadena, Calif , Sept 1, 1975
- 6 Klein, M J , and Stelzried, C T , "Calibration Radio Sources for Radio Astronomy Precision Flux-Density Measurements at 2295 MHz," *The Astronomical Journal*, Vol 81, 1078, 1976
- 7 Waltman, E , et al , "Compact Radio Sources at Declinations Greater Than 67 Degrees," submitted to *Astronomy and Astrophysics*, 1981
- 8 Fanselow, J L , private communication

Table 1 Northern polar survey results

(1) SOURCE NAME	(2) RIGHT ASCENSION				(3) DECLINATION				(4) REF CODE	(5) EXPERI- MENT	(6) TOTAL FLUX DENSITY (JY) +/-		(7) CORRELATED FLUX DENSITY (JY) +/-		(8) VISIBILITY +/-		(9) SPATIAL FREQUENCIES U V	
	HR	MIN	SEC	ERROR	DEG	MIN	SEC	ERROR										
0014+81	0	14	4 472	0 026	+81	18	28 72	0 28		B	1 0	0 3	0 34	0 03	0 3	0 1	-40 6	49 5
										C			0 35	0 03	0 4	0 1	-57 9	27 5
0016+73	0	16	54 200	0 022	+73	10	51 41	0 06	1	B	1 5	0 1	0 50	0 04	0 33	0 03	-42 9	46 6
										C			0 54	0 04	0 36	0 04	-49 2	40 3
0018+72	0	18	34 483	0 034	+72	56	3 93	0 30		B	0 8	0 2	0 11	0 01	0 14	0 04	-43 7	45 8
										C			0 11	0 01	0 14	0 04	-50 0	39 3
0027+70	0	27	17 026	0 033	+70	21	6 26	0 41		C	0 8	0 1	0 21	0 02	0 26	0 04	-49 8	39 1
0149+71	1	49	20 814	0 022	+71	0	20 76	0 30		B	1 5	0 2	0 18	0 02	0 12	0 02	-27 7	55 8
										C			0 16	0 01	0 11	0 02	-34 1	52 4
0153+74	1	53	4 339	0 019	+74	28	5 69	0 28		B	2 1	0 2	0 66	0 05	0 31	0 04	-28 3	56 3
										C			0 71	0 06	0 34	0 04	-34 5	52 9
0205+72	2	5	26 907	0 019	+72	15	16 40	0 28		B	0 8	0 2	0 41	0 03	0 5	0 1	-28 6	55 7
										C			0 38	0 03	0 5	0 1	-35 2	52 1
0212+73	2	12	49 935	0 019	+73	35	40 15	0 28		B	2 3	0 2	1 04	0 08	0 45	0 05	-28 6	56 0
										C			1 01	0 08	0 44	0 05	-34 8	52 6
0403+76	4	3	59 196	0 019	+76	48	52 75	0 31		B	4 4	0 1	0 11	0 01	0 025	0 002	-4 9	63 0
										C			0 14	0 01	0 032	0 002	-9 7	62 4
0454+84	4	54	57 213	0 019	+84	27	53 02	0 28		B	1 3	0 1	0 90	0 07	0 67	0 08	8 5	63 5
										C			0 80	0 06	0 62	0 07	4 2	63 9
0604+72	6	4	39 237	0 020	+72	49	27 21	0 29		B	1 0	0 2	0 19	0 02	0 19	0 04	20 2	59 1
										C			0 28	0 02	0 28	0 06	14 2	60 7
0615+82	6	15	32 80		+82	03	56 5		3	B	1 3	0 3	0 61	0 05	0 5	0 1	21 9	60 1
										C			0 64	0 05	0 5	0 1	20 0	60 7
0633+73	6	33	6 438	0 027	+73	27	35 75	0 41		C	0 9	0 3	0 36	0 03	0 4	0 1	20 2	59 2
0716+71	7	16	13 061	0 022	+71	26	14 92	0 41	1	B	0 7	0 3	0 11	0 01	0 16	0 07	31 8	53 8
										C			0 11	0 01	0 16	0 07	29 3	55 1
0718+79	7	18	8 892	0 019	+79	17	22 60	0 28		B	1 0	0 2	0 51	0 04	0 5	0 1	23 6	59 1
										C			0 50	0 04	0 5	0 1	28 2	57 2
0740+82	7	40	33 222	0 027	+82	49	24 17	0 29		A	1 6	0 3	0 42	0 04	0 26	0 06	60 4	21 5
										C			0 27	0 02	0 17	0 03	32 5	55 2
4C 71 07	8	36	21 358	0 009	+71	4	22 46	0 03	2	A	4 4	0 3	0 63	0 05	0 14	0 01	62 6	13 8
										A			0 37	0 03	0 084	0 009	13 3	60 3
0950+74	9	50	4 573	0 032	+74	50	7 67	0 29		A	1 2	0 3	0 19	0 02	0 16	0 04	63 8	-3 8
										A			0 12	0 01	0 10	0 03	54 2	33 8
1003+83	10	03	25 711	0 041	+83	04	56 52	0 40		A	1 1	0 3	0 43	0 04	0 4	0 1	63 1	11 5
1011+81	10	11	52 2		+81	20	16		3	A	0 6	0 3	<0 05		<0 08		63 1	11 4
										C			<0 05		<0 08		45 8	44 8
1039+81	10	39	27 821	0 024	+81	10	23 54	0 28		A			0 38	0 03			63 8	5 8
										C			0 63	0 05			48 8	41 5
1044+71	10	44	49 741	0 019	+71	59	26 73	0 28		A	1 0	0 2	0 73	0 06	0 7	0 2	63 8	-4 3
										C			0 95	0 08	1 0	0 2	56 0	30 6
1049+72	10	49	06 8		+72	15	42		3	A	0 8	0 2	<0 04		<0 05		62 6	-11 7
										C			<0 05		<0 06		58 1	26 7
1053+70	10	53	27 730	0 020	+70	27	47 78	0 28		A	0 6	0 3	0 59	0 05	1 0	0 5	62 0	-14 0
										C			0 66	0 05	1 1	0 6	59 4	23 7
1053+81	10	53	36 302	0 019	+81	30	35 41	0 23		A	0 6	0 3	0 42	0 04	0 7	0 4	63 9	3 9
										B			0 51	0 04	0 8	0 3	60 2	22 2
										C			0 50	0 04	0 8	0 4	51 0	38 8
1058+72	10	58	20 116	0 022	+72	41	44 72	0 28		A	1 0	0 2	0 56	0 05	0 6	0 1	63 6	-6 1
										C			0 21	0 02	0 21	0 05	56 7	29 4
1150+81	11	50	23 48		+81	15	10 2		3	A	1 2	0 3	0 76	0 07	0 6	0 2	63 2	-9 4
										C			0 88	0 07	0 7	0 2	58 0	27 4
1221+80	12	21	47 662	0 036	+80	56	41 03	0 28		A			0 18	0 02			61 9	-15 8
										C			0 21	0 02			60 5	21 4

Table 1 (contd)

(1) SOURCE NAME	(2) RIGHT ASCENSION				(3) DECLINATION				(4) REF CODE	(5) EXPERI- MENT	(6) TOTAL FLUX DENSITY (JY) +/-		(7) CORRELATED FLUX DENSITY (JY) +/-		(8) VISIBILITY +/-		(9) SPATIAL FREQUENCIES U V	
	HR	MIN	SEC	ERROR	DEG	MIN	SEC	ERROR										
1305+80	13	5	22 146	0 034	+80	24	21 27	0 24		A	0 7	0 3	0 15	0 02	0 21	0 10	58 5	-25 0
													0 13	0 01	0 19	0 08	-25 3	-57 4
													0 14	0 01	0 20	0 09	62 9	12 2
1345+73	13	45	14 2		+73	35	46		3	A	1 0	0 3	<0 04		<0 04		60 0	-20 6
										B			<0 05		<0 05		-24 1	-55 9
1357+76	13	57	42 177	0 035	+76	57	53 05	0 40		A			0 26	0 02			52 2	-35 3
1436+76	14	36	4 574	0 077	+76	18	23 82	0 45		A	1 3	0 3	0 07	0 01	0 05	0 01	51 4	-36 2
1448+76	14	48	56 496	0 022	+76	13	33 81	0 28		A	1 0	0 3	0 44	0 04	0 4	0 1	50 4	-37 5
3C 309 1	14	58	56 646	0 015	+71	52	11 15	0 06	2	B			0 37	0 03	0 4	0 1	-40 2	-47 6
										A			0 14	0 01			50 9	-35 8
										A			0 08	0 01			63 9	5 1
										B			0 53	0 05			-39 0	-47 2
1557+70	15	57	37 1		+70	49	45		3	B			0 14	0 01			-4 3	-59 6
										A	1 4	0 3	<0 05		<0 04		41 3	-45 1
										B			<0 05		<0 04		-49 4	-37 4
1612+79	16	12	21 8		+79	47	30		3	A	1 3	0 3	<0 04		<0 03		39 9	-48 6
1616+85	16	16	22 349	0 256	+85	9	26 04	0 47		B			<0 05		<0 04		-50 4	-38 2
										A	1 1	0 3	<0 05		<0 04		48 3	-41 6
1637+82	16	37	56 849	0 045	+82	38	18 48	0 30		B			0 07	0 01	0 06	0 02	-48 1	-41 9
										A	2 1	0 3	0 12	0 01	0 057	0 009	42 2	-47 3
1749+70	17	49	03 399	0 021	+70	06	39 44	0 25	1	B			0 24	0 02	0 11	0 02	-52 7	-35 7
										A	1 1	0 3	0 22	0 02	0 20	0 06	28 4	-52 8
1803+78	18	03	39 207	0 038	+78	27	54 08	0 23	1	B			0 34	0 03	0 31	0 09	-59 4	-21 4
										A	2 6	0 2	0 75	0 06	0 29	0 03	-59 2	-23 2
1807+69	18	7	18 503	0 020	+69	48	57 07	0 28		B	3 1	0 3	1 53	0 12	0 59	0 06	6 5	-61 7
1825+74	18	25	56 1		+74	19	05		3	B			1 01	0 08	0 33	0 04	7 2	-58 5
										A	1 0	0 3	<0 05		<0 05		27 6	-54 7
1928+73	19	28	49 333	0 040	+73	51	44 67	0 20	1	B			<0 05		<0 05		-60 7	-18 6
										A	3 0	0 3	0 45	0 04	0 15	0 02	13 0	-59 3
1946+70	19	46	12 034	0 028	+70	48	21 66	0 32		B			0 77	0 06	0 26	0 03	-63 9	-3 7
										A	0 7	0 3	0 17	0 02	0 2	0 1	10 2	-58 6
2007+77	20	07	20 478	0 060	+77	43	58 12	0 18	1	B			0 06	0 01	0 09	0 04	-64 0	- 3
										A	1 2	0 3	1 09	0 09	0 9	0 2	8 3	-61 3
2010+72	20	10	16 193	0 020	+72	20	20 81	0 28		B			0 68	0 05	0 6	0 1	-64 0	5
										A	1 5	0 3	0 59	0 05	0 39	0 09	5 1	-59 8
2136+82	21	36	2 188	0 029	+82	25	38 61	0 30		B			0 53	0 04	0 35	0 08	-64 0	3 9
										A	1 1	0 3	0 17	0 02	0 15	0 05	-10 7	-62 2
2229+69	22	29	11 643	0 019	+69	31	2 79	0 29		B			0 12	0 01	0 11	0 03	-61 4	18 6
										A			0 43	0 04			-23 3	-54 7
										B			0 54	0 04			-56 1	30 1

Table 2. Notes for Table 1

Column	Description								
1	Source name								
2,3	Source position right ascension and declination The positions are referenced to the equinox of 1950 0, and elliptical aberration terms are included so as to agree with past astronomical convention								
4	Source position reference code If no reference code is given, the position was estimated from the delay and delay rate observables Otherwise the position was obtained from one of the following references								
	<table> <tr> <th>Code</th><th>Reference</th></tr> <tr> <td>1</td><td>Waltman et al (Ref 7)</td></tr> <tr> <td>2</td><td>JPL Reference Frame Catalog (Ref 8)</td></tr> <tr> <td>3</td><td>Bonn Catalog (Ref 2)</td></tr> </table>	Code	Reference	1	Waltman et al (Ref 7)	2	JPL Reference Frame Catalog (Ref 8)	3	Bonn Catalog (Ref 2)
Code	Reference								
1	Waltman et al (Ref 7)								
2	JPL Reference Frame Catalog (Ref 8)								
3	Bonn Catalog (Ref 2)								
	The detected sources whose positions came from these references had well-known positions and allowed instrumental biases to be calibrated								
5	Experiment codes A = 3 March 1980 B = 12/13 March 1980 C = 28 March 1980								
6	Total flux density and uncertainty in Jy								
7	Correlated flux density and uncertainty in Jy								
8	Fringe visibility and uncertainty Defined as the ratio of correlated flux density to total flux density or the fraction of flux density detected by the interferometer as coming from the milliarcsecond core								
9	Spatial frequencies U is the east-west component of the baseline projection against the sky, and V is the north-south component of the baseline projection against the sky These are given in millions of wavelengths								

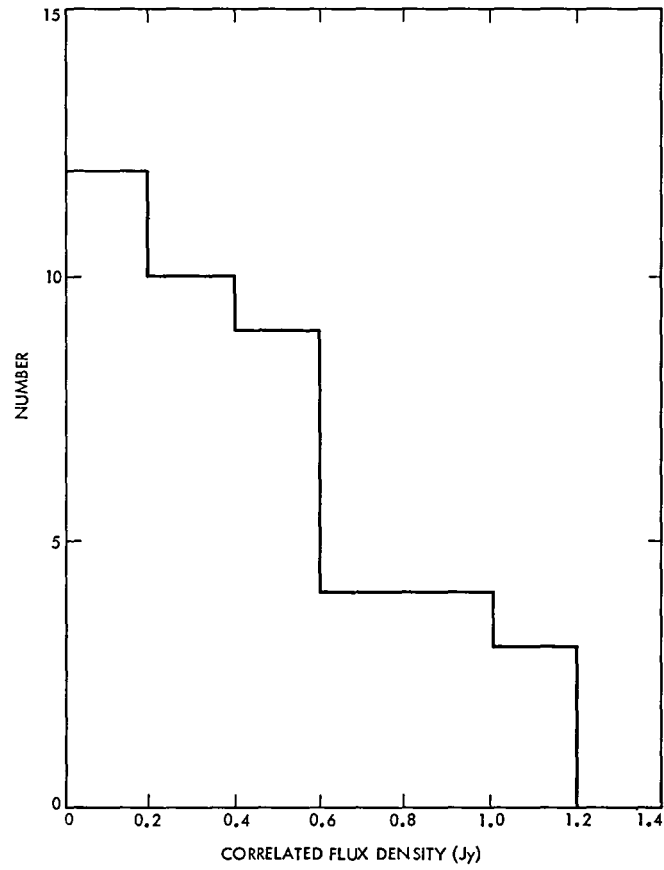


Fig. 1. Correlated flux density histogram

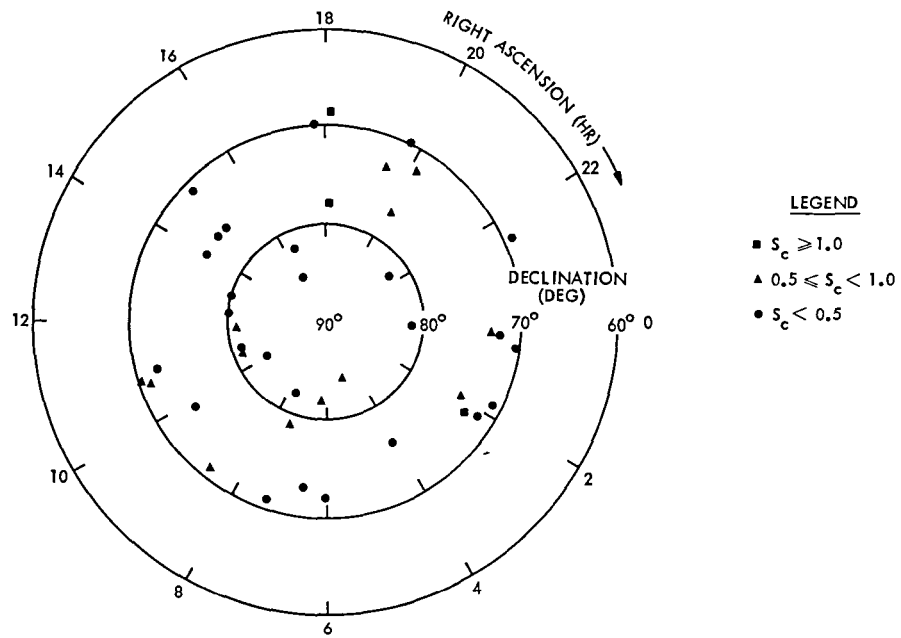


Fig. 2. Sky distribution plot of detected sources

Arcsecond Positions for Milliarcsecond VLBI Nuclei of Extragalactic Radio Sources. Part I: 546 Sources

D. D. Morabito, R. A. Preston, and M. A. Slade
Tracking Systems and Applications Section

D. L. Jauncey
CSIRO, Division of Radiophysics,
Sydney, NSW, Australia

VLBI measurements at 2290 MHz on baselines of $\sim 10^4$ km between Deep Space Network stations have been used to identify sources that are suitable for precision astrometric applications. In addition, the observations are used to determine the positions of the milliarcsecond nuclei in 546 extragalactic radio sources. Estimated accuracies generally range from $\sim 0''.5$ to $1''.0$ in both right ascension and declination, with $\sim 6\%$ of the sources having estimated uncertainties $\geq 3''.5$ in at least one position component. Arcsecond positions serve as a useful starting point in the construction of high-precision VLBI reference frames, and are also important for unambiguous determination of optical counterparts to compact radio sources.

I. Introduction

Two of the quantities that may be measured by very long baseline interferometry (VLBI) are time delay and fringe frequency. Time delay is defined as the difference in time of arrival of signals from a celestial radio source at two antennas plus an instrumental delay. Fringe frequency is defined as the difference in the doppler shift of the received signals as induced by Earth rotation plus the local oscillator frequency offset between the two antennas. A measurement of time delay places the position of an observed source along a particular arc in the sky if the contribution due to instrumental delay can be removed. Correspondingly, a measurement of fringe frequency places the position of the source along a generally different sky arc if the contribution due to instrumental local oscillator frequency offset can be removed. These arguments assume that baseline coordinates and atmospheric effects are known. Observations of sources with accurately known positions allow instrumental delays and frequency offsets to be determined. Hence, a measurement of time delay and fringe

frequency can be used to uniquely determine the right ascension and declination of a source.

This technique was first described by Cohen and Shaffer (1971) (Ref. 1), who determined the positions of 35 compact sources with typical accuracies of $1''$ to $3''$. More accurate positions for $\sim 10^2$ compact sources have been determined with the VLBI technique of "bandwidth synthesis," which provides extremely accurate measurements of time delay (Refs. 2, 3). In this article we use relatively crude time delay measurements along with fringe frequency measurements to determine the positions of the milliarcsecond nuclei in 546 extragalactic sources with accuracies of generally $\sim 1''$.

The observations were not performed primarily for position measurements, but were part of the development of an all-sky VLBI catalog of milliarcsecond radio sources at 2290 MHz sponsored by the Deep Space Network Advanced Systems Program (Ref. 4). The catalog is being formed by searching for

VLBI components in known extragalactic radio sources. The catalog will eventually contain about 10^3 sources and will be useful in the creation of high-precision celestial reference frames and in investigations into the nature of compact extragalactic radio sources. We expect to determine arcsecond positions for a large majority of those sources, which will constitute at least an order-of-magnitude improvement in positional accuracy for most of the sources.

Arcsecond positions serve as a useful starting point in the construction of high-precision VLBI reference frames. VLBI reference frames are used for Δ VLBI navigation (Ref. 5), clock synchronization (Ref. 6) and baseline determination between widely separated Deep Space Stations (Refs. 7, 8) and measurement of Earth rotation orientation (Ref. 9). The results of this article are presently being utilized as a first step in the formation of a VLBI reference frame of about 100-200 sources in which relative source positions will probably be determined to $\sim 10^{-3}$ " by means of high-accuracy bandwidth synthesis time delay measurements (Refs. 2, 10). The arcsecond measurements are also being used for unambiguous determination of optical counterparts to the compact radio sources and for identifying the location of the compact source in more extended optical or radio objects.

II. The Observations

The observations were performed with pairs of antennas within NASA's Deep Space Network on either California-Australia or California-Spain baselines (see Table 1) in 30 separate observing sessions between 1974 and 1978. A list of experiments appears in Table 2. The observations were performed at 2290 MHz with right circular polarization. The receiver chain generally consisted of an S-band traveling wave maser, followed by a special phase stable S-band VLBI receiver which converted the RF signal to an IF of 50 MHz. The NRAO Mark II VLBI recording system then recorded a 18-MHz data bandwidth by digitally sampling at a 4-Mbs rate (Ref. 11). Digital sampling and phase stability of the receiver chain were controlled by rubidium or hydrogen maser atomic clocks. Each source was observed for a few minute interval with most sources being observed only once.

III. Method of Position Determination

A. The Measurement of Time Delay and Fringe Frequency

Values of time delay and fringe frequency were obtained with the NRAO Mark II VLBI cross-correlation processor and postprocessor programs. Each observation was broken into a

number of 30-second segments. For each segment residual values were calculated between measured values of time delay and fringe frequency and premeasurement estimates of these quantities. All usable segments of each observation were then averaged to produce mean values of time delay and fringe frequency residuals and corresponding rms scatters. Typical values for the rms time delay residual scatter were 30 nsec and for the rms fringe frequency residual scatter, ~ 2 mHz. These scatters were assumed to be estimates of random measurement error unless the values appeared unrealistically low, in which case default values of 20 nsec and 2 mHz were used.

B. Theoretical Models of Time Delay and Fringe Frequency

To estimate source positions, we must have theoretical models for the quantities of time delay (τ) and fringe frequency (ν). Such models may be expressed by

$$\tau = \tau_g + \tau_t + \tau_i$$

$$\nu = \nu_g + \nu_t + \nu_i$$

where τ_g and ν_g are the geometric contributions, τ_t and ν_t are the tropospheric contributions, and τ_i and ν_i are the instrumental contributions. No ionospheric contributions were considered since (1) the contribution of the unmodeled ionosphere is generally no larger than the assumed random measurement errors in time delay and fringe frequency, and (2) ionospheric contributions are difficult to estimate.

The geometric contribution to time delay is merely the dot product of a unit vector in the source direction and the baseline vector in light seconds, corrected to account for the Earth not being in the same rotational position when a particular wavefront arrives at each of the two antennas (Ref. 1). The geometric contribution to fringe frequency is merely the time derivative of τ_g (in Hz). The uncertainty in the baseline vector coordinates between Deep Space Network antennas is ~ 1 m (3 nsec), which is negligible compared to the assumed random measurement errors in time delay and fringe frequency.

The tropospheric contribution to time delay was determined by calculating the effective tropospheric signal delay at each antenna according to a cosecant elevation angle dependence and then differencing the resultant values. The zenith values of the dry and wet components of the tropospheric delay were assumed to be 6.7 and 0.5 nsec, respectively. Errors in the troposphere model should amount to ≤ 1 nsec in zenith path length delay and when mapped to the observed elevation angles are insignificant compared to the assumed random measurement errors in time delay and fringe frequency.

C. Calibration of Instrumental Delays and Frequency Offsets

Observations of sources with well-known positions allowed instrumental delays and frequency offsets to be determined. For each experiment, such calibration sources were scattered in time among the sources whose positions we wished to determine. A list of the 82 calibration sources utilized appears in Table 3. References from which the calibration source positions were obtained appear in Table 4. Most of the calibration source positions are VLBI determined and can be referenced to the FK4 reference frame (Ref. 12) with an accuracy of $\sim 0''.1$ (Ref. 10). However, a few calibration source positions are optically determined and may have uncertainties as high as $1''.0$ in the FK4 reference frame.

At least four position calibration sources were employed per experiment. Table 2 displays the number of calibration sources used in each experiment as well as the number of source positions determined. Within an individual experiment, the time delay and fringe frequency residuals of calibration sources often exhibited nearly linear drifts with time over the experiment. For each experiment, a straight line was fit in a weighted least-mean-square sense through the time delay residuals of the calibration sources. This resulted in estimates of instrumental time delay and instrumental time delay rate. In a similar manner, a straight line was fit to the fringe frequency residuals of the calibration sources, resulting in estimates of instrumental frequency offset and instrumental frequency offset rate. In each experiment, the scatters of the delay and fringe frequency data points about the straight line fits are consistent with the estimated data uncertainties.

D. Position Determination

For each observation of a noncalibration source, residuals were formed between the measured values of time delay and fringe frequency and the corresponding values predicted by the best premeasurement models for these quantities. After these residuals were corrected for instrumental effects (see III-C), a least-mean-square differential correction scheme was used to estimate corrections to source position in right ascension and declination as well as the uncertainties associated with these corrections. In this calculation of the position uncertainty estimates, only the random errors in the measurement of time delay and fringe frequency were considered. To account for unmodeled effects, $0''.5$ was added in a rss manner to the formal estimates of right ascension and declination uncertainty for each experiment but one (75-06-17), to which we added $1''.0$. Since the position correction vector was calculated in coordinates precessed to the date of the experiment, the resultant new source position was subsequently precessed back to 1950.0 coordinates.

IV. Results

The calculated positions of 546 sources and the corresponding uncertainties are shown in Table 5. The positions are referenced to the equinox of 1950.0 and elliptical aberration terms are included so as to agree with past astronomical convention. The source positions and position uncertainties for sources which were multiply observed were estimated from a weighted average.

Seventy sources were observed two or more times, and the scatters in their position estimates are consistent with the estimated uncertainties. Figure 1 displays a histogram of the ratio of the weighted rms scatter of the individual position coordinate estimates from the mean position of each multiply observed source to the weighted rms estimate of position coordinate uncertainty. For the 70 multiply observed sources, the rms value of this ratio is ~ 0.6 for both right ascension and declination. In addition, the individual ratio values never exceed 1.6.

For 35 sources, we could compare our position estimates with higher accuracy position estimates (Refs. 10, 13, 14). Figure 2 displays the ratio of the absolute value of the difference between our source position estimate and the more accurate value to our estimate of position uncertainty. For these 35 sources, the rms value of this ratio is ~ 1.3 for both right ascension and declination, but declines to ~ 0.9 if we eliminate one source (P2145+06) whose ratios are anomalously high. For both right ascension and declination, the value of the ratio exceeds 2.0 for only 3 sources, reaching a maximum value of 5.6. Hence, these position estimate comparisons along with the multiple observation comparisons indicate our position uncertainty estimates are realistic.

Figures 3(a) and 3(b) show histograms of the number of sources versus estimated position uncertainty for declination and right ascension, respectively. Estimated accuracies generally range from $\sim 0''.5$ to $1''.0$ in both right ascension and declination, but $\sim 6\%$ of the sources have uncertainties $\geq 3''.5$ in at least one position component. Most of the sources with high uncertainties were near the equator where the position-measuring technique can have lower sensitivity (see Ref. 1).

V. Summary

Positions for the milliarcsecond VLBI nuclei of 546 extragalactic sources have been determined to an accuracy of $\sim 1''$. The reliability of the determined positions has been demonstrated by testing the repeatability of multiple observations on the same source and by position comparisons with more accurate radio catalogs.

Acknowledgments

We appreciate the assistance of G Trager in data processing and L J Skjerve, D J Spitzmesser, and the personnel of the Deep Space Network for aid in performing the observations

References

- 1 Cohen, M H , and Shaffer, D B , *Astron J* , 76, 91, 1971
- 2 Fanselow, J L , Thomas, J B , Cohen, E J , MacDoran, P F , Melbourne, W G , Mulhall, B D , Purcell, G H , Rogstad, D H , Skjerve, L J , Spitzmesser, D J , Urech, J , and Nicolson, G , *Time and the Earth's Rotation*, IAU Symp 82, D Reidel, p 199, 1979
- 3 Clark, T A , Hutton, L K , Marandino, G E , Counselman, C C , Robertson, D J , Shapiro, I I , Wittels, J J , Hinteregger, H F , Knight, C A , Rogers, A E E , Whitney, A R , Niell, A E , Ronnang, B U , and Rydbeck, O E H , 1976, *Astron J* , 81, p 599
- 4 Preston, R A , Morabito, D D , Williams, J G , Slade, M A , Harris, A W , Finley, S G , Skjerve, L J , Tanida, L , Spitzmesser, D J , Johnson, B , Jauncey, D L , Bailey, A , Denise, R , Dickenson, J , Livermore, R , Papij, A , Robinson, A , Taylor, C , Alcazar, F , Luaces, B , and Munoz, D , "Establishing a Celestial VLBI Reference Frame — I Searching for VLBI Sources," *DSN Progress Report 42-46*, Jet Propulsion Laboratory, Pasadena, Calif , Aug 15, 1978
- 5 Brunn, D L , Preston, R A , Wu, S C , Siegel, H L , and Brown, D S , "VLBI Spacecraft Tracking System Demonstration Part I, Design and Planning," *DSN Progress Report 42-45*, Jet Propulsion Laboratory, Pasadena, Calif , June 15, 1978
- 6 Mulhall, B D L , "DSN VLBI System Status and Plans," *DSN Progress Report 42-46*, Jet Propulsion Laboratory, Pasadena, Calif , Oct 15, 1978
- 7 MacDoran, P F , "Radio Interferometry for International Study of the Earthquake Mechanism," *Acta Astronautica*, Vol 1, pp 1427-1444, 1974
- 8 Niell, A E , Ong, K M , MacDoran, P F , Resch, G M , Morabito, D D , Claflin, E S , and Dracup, J F , "Comparison of a Radio Interferometric Differential Baseline Measurement with Conventional Geodesy," *Tectonophysics*, 52, pp 49-58, 1979
9. Shapiro, I I , et al , "Transcontinental Baselines and the Rotation of the Earth Measured by Radio Interferometry," *Science*, 186, pp 920-922, 1974
- 10 Fanselow, J L , Sovers, O J , Thomas, J B , Bletzacker, F R , Kearns, T J , Cohen, E J , Purcell, G H , Jr , Rogstad, D H , Skjerve, L J , and Young, L E , *Reference Coordinate Systems for Earth Dynamics*, pp 351-357, D Reidel, 1981
- 11 Clark, B G , *Proc of the I E E E* , 61, p 1242, 1973
- 12 Fricke, W , and Kopff, A , *Veroeff Astron* , Rechen-Institut, Heidelberg, No 10, 1963
13. Elsmore, B , and Ryle, M , *MNRAS* , 174, p 411, 1976
- 14 Wade, C M , and Johnston, K J , *Astron J* , 82, p 791, 1977

Table 1. Participating Deep Space Station (DSS) observatories

Location	Designation	Diameter, m	Baseline Length	
			Kilometers	Wavelengths
Tidbinbilla, Australia	DSS 42	26	10.6 × 10 ³	8.1 × 10 ⁷
	DSS 43	64		
Goldstone, California	DSS 13	26	8.4 × 10 ³	6.4 × 10 ⁷
	DSS 14	64		
Madrid, Spain	DSS 61	26	8.4 × 10 ³	6.4 × 10 ⁷
	DSS 63	64		

Table 2. Equipment list

Experiment			DSS stations	Number of position calibrator sources	Number of source positions determined
Yr	Mo	Day			
74	07	31	14/42	6	6
74	08	13	14/42	4	11
75	06	17	14/42	4	19
75	08	24	14/42	4	15
75	09	15	14/42	6	17
75	09	21	13/63	8	27
76	10	26A	13/43	9	21
76	10	26B	13/43	6	20
76	11	09	13/63	4	8
76	11	11	13/63	9	22
76	11	14	13/43	5	28
77	02	12	13/43	5	8
77	02	25	13/43	6	26
77	04	21	13/43	4	14
77	04	22	13/43	5	18
77	06	15	13/43	4	18
77	09	11	14/61	4	8
77	09	28	13/63	7	48
77	10	11	13/63	8	20
77	10	27	13/43	5	16
77	11	01	13/43	4	27
77	11	21	13/43	6	20
77	12	11	13/43	4	21
77	12	13	13/43	5	28
78	01	09	13/43	5	57
78	02	20	13/43	5	21
78	03	14	13/63	4	24
78	03	21	13/63	4	37
78	04	04	13/63	4	32
78	04	16	13/43	5	11

Table 3. List of calibration sources

Source	Position reference	Source	Position reference	Source	Position reference
P 0104-408	1	4C 55 16	1	NRAO 512	1
P 0106+01	1	OJ 287	1	3C 345	1
P 0113-118	1	B2 0912+29	13	P 1645+17	2
DA 55	1	4C 39 25	1	DW 1656+05	1
3C 48	13	AO 0952+17	1	NRAO 530	1
DW 0224+67	1	OK 290	13	P 1741-038	1
GC 0235+16	1	3C 232	13	3C 371	1
3C 84	1	GC 1004+14	6	3C 395	13
P 0332-403	1	P 1104-445	1	OV-236	1
NRAO 140	1	P 1116+12	5	P 1933-400	1
CTA 26	1	P 1127-14	1	OV-198	1
NRAO 150	1	B2 1215+30	13	P 2012-017	11
P 0402-362	1	ON 231	1	P 2021-330	12
P 0420-01	1	3C 273	1	3C 418	13
P 0428+20	13	3C 274	2	GC 2047+09	10
3C 119	13	3C 275 1	7		
3C 120	1	3C 279	1	P 2134+004	1
P 0438-43	1	3C 286	13	P 2137+20	10
NRAO 190	1	P 1354+19	8	P 2145+06	1
3C 138	13	OQ 208	13	VRO 42 22 01	1
3C 147	13	OQ 172	13	CTA 102	1
DA 193	1	OR 103	1	OY-172 6	1
P 0605-08	1	P 1510-08	1	P2245-328	1
P 0607-15	1	GC 1514+19	13	3C 454.3	1
P 0736+01	3	P 1514-24	9	GC 2254+07	13
OI 363	1	P 1546+027	3	P 2319+07	4
DW 0742+10	1	DW 1555+00	1	P 2345-16	1
OJ 425	1	DA 406	1		

Table 4 References for calibration sources

- 1 Fanselow, J L , Sovers, O J., Thomas, J. B , Bletzacker, F R , Kearns, T J , Cohen, E J , Purcell, G H Jr , Rogstad, D H , Skjerve, L J , Young, L E , *Reference Coordinate Systems for Earth Dynamics*, 351-357, D Reidel Publishing Company, 1981
- 2 Burbidge, G R , and Crowne, A H , *Ap J Suppl Series*, 40, p 583, 1979
- 3 Burbidge, G R , Crowne, A H , and Smith, H E , *Ap J. Suppl Series*, 33, p 113, 1977.
- 4 Veron, M. P , and Veron, P., *Astr Ap* , 42, p 1, 1975
- 5 Kristian, J and Sandage, A , *Ap J* , 162, 391, 1970
- 6 Wills, D , and Wills, B J , *Ap J Suppl Series*, 31, p 143, 1976.
- 7 Adgie, R L , Crowther, F H , and Gent, H , *MNRAS* 159, p 233, 1972
- 8 Veron, M P , *Astr Ap* , 20, p 471, 1972
- 9 Hunstead, R W , *Nature*, 233, p 401, 1971
- 10 Hoskins, D , Murdoch, H S , Adgie, R L , Crowther, F. H , and Gent H , *MNRAS* , 166, p 235, 1979.
- 11 Wall, J L , *Ap J (Letters)*, 15, p 101, 1973
- 12 Veron, M P , and Veron, P., *Astr Ap Suppl* , 29, p 149, 1977
- 13 Fanselow, J. L , private communication, 1980

Table 5 Source positions

SOURCE NAME	HR	RIGHT MIN	ASCENSION SEC	ERROR	DEG	DECLINATION MIN	SEC	ERROR	SOURCE NAME	HR	RIGHT MIN	ASCENSION SEC	ERROR	DEG	DECLINATION MIN	SEC	ERROR
GC 0003+38	00	03	22.368	0 035	+38	03	33 37	0 52	OD 120	02	11	59 754	0 046	+17	08	51 86	0 91
NRAO 5	00	03	40.241	0 081	-06	40	16 63	1 15	P 0216+011	02	16	32 566	0 068	+01	07	11 71	1 15
P 0005-239	00	05	27 468	0.033	-23	55	59 91	0 46	P 0217-189	02	17	00 337	0 039	-18	56	24 57	0.57
P 0005-262	00	05	53 498	0 045	-26	15	53 14	0 65	3C 66A	02	19	29 995	0 045	+42	48	29 65	0 63
III ZW 2	00	07	57 865	2 108	+10	41	30 48	34 51	P 0219-164	02	19	38 310	0 037	-16	28	55 46	0.56
GC 0007+17	00	07	59 377	0 035	+17	07	37 75	0.54	3C 66B	02	20	01 732	0 056	+42	45	54 59	0 87
P 0008-264	00	08	28 858	0 036	-26	29	14 58	0 50	GC 0221+06	02	21	50 047	0 051	+06	45	49 17	0 85
GC 0010+40	00	10	54 332	0 042	+40	34	56 75	0 69	P 0229+13	02	29	02 532	0 038	+13	09	40 70	0 67
P 0011-046	00	11	20.643	0 102	-04	40	32 65	1 39	P 0234-301	02	34	21 633	0 047	-30	06	56 09	0.64
GC 0012+31	00	12	29 879	0 037	+31	59	32 94	0 55	CTD 20	02	34	55 593	0 034	+28	35	11 28	0 50
GC 0019+05	00	19	58.289	0 220	+05	51	22 30	3 73	GC 0235+16	02	35	52 647	0 025	+16	24	03 69	0.38
P 0019+058	00	19	58.291	0 129	+05	51	22 17	2 31	GT 0236+61	02	36	53 910	0 218	+61	01	39 13	2 21
OB 337 7	00	22	46 685	0 046	+39	02	58 58	0 71	P 0237-027	02	37	13 721	0 176	-02	47	32 84	2 37
OB 338	00	24	02 809	0.036	+34	52	06 29	0 53	GC 0237+04	02	37	14 427	0 067	+04	03	29 31	1.00
OB 343	00	26	34 849	0 104	+34	39	57 97	2 02	P 0237-23	02	37	52 787	0 047	-23	22	06 45	0 61
P 0027+056	00	27	11 367	0 098	+05	38	05 40	1 80	P 0238-084	02	38	37 382	0 056	-08	28	09 38	0 77
P 0035+23	00	35	19.774	0 047	+23	50	42 43	0 62	OD 166	02	39	47 109	0 040	+10	48	16 22	0 63
GC 0035+12	00	35	41 931	0 045	+12	11	01 69	0 68	P 0240-217	02	40	19 234	0 402	-21	45	09 63	3.57
P 0047+023	00	47	08 930	0 060	+02	20	43 11	1 33	GC 0250+17	02	50	46 337	0 038	+17	53	29 97	0 56
P 0048-09	00	48	09.003	0 925	-09	45	08 94	14 36	OD 094 7	02	56	46 979	0 066	+07	35	45 50	0 86
P 0048-071	00	48	36 190	0 061	-07	06	20 65	0 91	OE 400	03	00	10 114	0 034	+47	04	33 64	0 51
P 0048-427	00	48	48.972	0 067	-42	47	52 15	0 83	GC 0301+33	03	01	35 678	0 039	+33	37	04 74	0.66
DW 0055+30	00	55	05 632	0.033	+30	04	56 66	0 51	OE 110	03	06	20 942	0 048	+10	17	51 56	0 78
P 0055-059	00	55	32.966	0 045	-05	56	03 66	0 71	MJ 0309+41	03	09	44 793	0 038	+41	08	48 67	0 59
P 0056-00	00	56	31.839	0 069	-00	10	08 84	21 28	P 0312+10	03	12	38 352	0 197	+10	01	40 84	2.20
0104+321	01	04	39.247	0 053	+32	08	44 33	0 87	CTA 21	03	16	09 132	0 050	+16	17	40 20	0.86
P 0106+01	01	06	04.589	0 031	+01	19	00 04	0 61	3C 84	03	16	29 567	0 036	+41	19	51 86	0 52
P 0108-079	01	08	19 018	0 046	-07	57	37 67	0 70	P 0317+188	03	17	00 064	0 041	+18	50	42 10	0 58
GC 0108+38	01	08	47 275	0 045	+38	50	32 78	0 65	P 0319+12	03	19	08 153	0 086	+12	10	32 31	0 80
GC 0109+22	01	09	23 595	0 043	+22	28	44 24	0 60	P 0327-241	03	27	43 892	0 036	-24	07	23 15	0 54
GC 0110+31	01	10	03 792	0 090	+31	52	24 02	2 35	P 0329-255	03	29	00 470	0 043	-25	34	53 23	0 70
P 0111+021	01	11	08 614	0 058	+02	06	23 99	1 05	P 0346-163	03	46	21 706	0 068	-16	19	23 84	0 80
P 0112-017	01	12	43 935	0 045	-01	42	55 24	0 77	P 0348-120	03	48	48 961	0 249	-12	02	18 08	3.45
P 0115+02	01	15	43 683	0 036	+02	42	18 76	0 66	CTD 26	04	00	03 589	0 036	+25	51	46 55	0.53
GC 0116+08	01	16	24 203	0 098	+03	14	05 26	5 67	P 0400-319	04	00	23 602	0 035	-31	57	42 04	0.52
P 0118-272	01	18	09.508	0 043	-27	17	07 38	0 58	P 0403-13	04	03	13 983	0 130	-13	16	18 17	1.55
P 0119+11	01	19	03 134	0 042	+11	34	08 51	0 68	P 0405-385	04	05	12 020	0 037	-38	34	26 07	0.53
GC 0119+04	01	19	21.482	0 034	+04	06	42 62	0 52	P 0405-12	04	05	27 445	0 052	-12	19	32 44	0 67
GC 0119+24	01	19	54 290	0 039	+24	46	51 68	0 73	P 0405-331	04	05	38 518	0 087	-33	11	42 40	0 83
P 0122-00	01	22	55.304	0 035	-00	21	33 58	0 55	GC 0406+12	04	06	35 472	0 037	+12	09	49 28	0 57
P 0123+25	01	23	57 269	0.038	+25	43	27 49	0 65	P 0406-127	04	06	45 218	0 091	-12	46	37 39	1.15
3C 43	01	27	15 017	0.033	+23	22	52 34	0 50	3C 109	04	10	54.890	0 047	+11	04	40 32	0 65
P 0130-17	01	30	17 637	0 067	-17	10	11 21	0 71	P 0413-21	04	13	53 539	0 046	-21	03	49 86	0 59
OC-259	01	35	17 119	0 046	-24	46	08 92	0 64	P 0414-189	04	14	23 323	0 052	-18	58	29 50	0 59
P 0136+176	01	36	59 331	0 039	+17	37	56 03	0 58	3C 111	04	15	00 602	0 035	+37	54	19 52	0.54
P 0139-09	01	38	55 651	1 515	-09	43	32 68	23 65	P 0420+022	04	20	16 107	0 053	+02	12	28 83	1 10
P 0142-278	01	42	44 980	0 047	-27	48	35 53	0 62	OF 036	04	21	32 765	0 113	+01	57	31 63	2.14
GC 0144+20	01	44	13 602	0 039	+20	55	26 58	0 58	P 0422+00	04	22	12 516	0 041	+00	29	16 65	0 70
OC 079	01	46	45 625	0 038	+05	40	59 27	0 58	P 0423+051	04	23	57 289	0 042	+05	11	36 40	0 72
GC 0148+27	01	48	37 349	0 042	+27	29	49 99	0 83	P 0425+048	04	25	08 626	0 062	+04	50	29 55	1 33
P 0150-334	01	50	56 961	0 041	-33	25	10 81	0 55	P 0426-380	04	26	54 701	0 035	-38	02	52 21	0 51
P 0156-14	01	56	18 093	0 233	-14	27	36 50	3 33	P 0434-188	04	34	48 942	0 048	-18	50	47 96	0 58
3C 57	01	59	30.304	0 117	-11	46	59 06	1 44	DW 0436+50	04	36	49 749	0 219	+50	22	20 48	2.52
P 0201+113	02	01	06.076	0 055	+11	20	21 98	0 82	P 0439-337	04	39	41 941	0 042	-33	45	44 06	0 55
P 0202+14	02	02	07 449	0 036	+14	59	50 12	0 57	GC 0444+63	04	44	42 345	0 056	+63	26	55 93	0 65
DW 0202+31	02	02	09 663	0 037	+31	58	10 25	0 55	P 0446+11	04	46	21 222	0 053	+11	16	17 61	0 84
P 0202-17	02	02	34 481	0 046	-17	15	39 24	0 59	P 0451-28	04	51	15 084	0 044	-28	12	28 80	0 60

Table 5 (contd)

SOURCE NAME	HR	RIGHT ASCENSION MIN SEC	ERROR	DEG	DECLINATION MIN SEC	ERROR	SOURCE NAME	HR	RIGHT ASCENSION MIN SEC	ERROR	DEG	DECLINATION MIN SEC	ERROR	
P 0454+06	04	54	26 459	0 043	+06	40 29 35	0 74	OJ 320	08	12 10 708	0 041	+36	44 27 46	0 66
OF-292	04	54	57 299	0 036	-23	29 28 28	0 53	P 0812+02	08	12 47 287	0 040	+02	04 12 76	0 76
P 0456+060	04	56	08 200	0.054	+06	03 33 15	0 98	3C 197 1	08	17 59 852	0.116	+47	12 19 95	1.41
P 0457+024	04	57	15 549	0 049	+02	25 05 45	1 06	P 0820+22	08	20 28 601	0 081	+22	32 43 60	1.32
P 0458-02	04	58	41.353	0 043	-02	03 34 19	0 76	GC 0820+29	08	20 36 578	0 052	+29	38 10 61	0 95
GC 0459+06	04	59	34 814	0 039	+06	04 51 42	0 70	GC 0820+56	08	20 53 196	0.044	+56	02 27 44	0 58
OG 003	05	00	45 184	0 039	+01	53 53.80	0 65	GC 0821+39	08	21 37 307	0.037	+39	26 28 24	0 54
P 0511-220	05	11	41 822	0 037	-22	02 41.38	0 54	P 0823+033	08	23 13 566	0.025	+03	19 15 07	0.59
0528-250	05	28	05.156	0.049	-25	05 44 19	0 55	4C 35 20	08	24 26 606	0 044	+35	35 00 75	0 76
OG-247	05	28	05 215	0 037	-25	05 44 65	0 54	G 256 6+0 6	08	26 12.011	0 079	-37	21 06.28	0 86
P 0528+134	05	28	06.755	0 036	+13	29 42 29	0 56	B2 0827+24	08	27 54 397	0 040	+24	21 07 52	0 69
GC 0537+53	05	37	13 514	0 046	+53	10 54.20	0 58	OJ 448	08	28 48 014	0.034	+49	23 32 65	0 51
GC 0554+58	05	54	52 468	0 039	+58	03 50.87	0 59	MA 0829+04	08	29 10 927	0.044	+04	39 50 40	2 83
P 0555-132	05	55	44 054	0 361	-13	17 46.55	5.37	OJ 451	08	30 31.977	0.037	+42	34 19.49	0 57
3C 151	06	00	56 855	0 039	+44	14 10 76	0 61	4C 55 16	08	31 04 400	0 046	+55	44 41.33	0 56
GC 0601+57	06	01	22 080	0 049	+57	53 31 97	0 69	GC 0833+58	08	33 23 779	0 049	+58	35 30.19	0 61
OH 404.1	06	02	20 043	0.035	+10	30 25 69	0 54	GC 0834+25	08	34 42 304	0 038	+25	04 54 18	0 71
GC 0602+67	06	02	41 842	0.357	+57	20 50 18	2.95	4C 71 07	08	36 21 630	0 046	+71	04 21 92	0 53
P 0606-223	06	06	53.379	0 048	-22	19 46 31	0 62	3C 207	08	38 01.729	0 051	+13	23 05 71	0 84
GC 0611+48	06	11	15.403	0 041	+48	20 02.25	0 63	GC 0839+18	08	39 14 177	0 068	+18	46 25 66	1 03
OH 119 5	06	11	31 724	0 310	+13	55 06 70	4 56	GC 0850+53	08	50 50 150	0 047	+58	08 55 73	0 57
GC 0636+68	06	36	47 653	0 047	+68	01 27.14	0 56	OJ 287	08	51 57 311	0 070	+20	17 57 34	1 04
P 0637-337	06	37	31 171	0 045	-33	43 12.82	0 60	CC 0854+21	08	54 04 725	0 066	+21	23 17 83	1 22
OH 368.8	06	41	26.319	0 043	+39	17 55 12	0 68	3C 212	08	55 55 533	0 052	+14	21 24 75	1.09
OH 471	06	42	53 070	0 037	+44	54 30.23	0 56	OJ 499	08	59 40.021	0 039	+47	02 56.72	0 56
OH 577 1	06	46	04.123	0 049	+60	05 14.13	0 61	P 0859-14	08	59 55 059	0 085	-14	03 39.93	1 15
GC 0650+37	06	50	35 278	0 042	+37	09 27.21	0 63	GC 0900+42	09	00 58 735	0 044	+42	50 01 42	0 57
OH-090	06	53	41 212	0 041	-03	19 11.68	0 78	GC 0902+49	09	02 00 394	0 040	+49	02 49 46	0 60
GC 0707+47	07	07	02 539	0 044	+47	37 07 92	0 59	3C 216	09	06 17 253	0.031	+43	05 59 12	0 43
OI 417	07	10	03 990	0 050	+43	54 28.65	0 70	P 0906+01	09	06 35 308	0 080	+01	33 46.53	1.55
OI 318	07	11	05 643	0 034	+35	39 51.90	0 52	GC 0913+39	09	13 39 507	0 045	+39	07 02 28	0 59
3C 179	07	23	04 272	0 042	+67	54 52 87	0 53	OK 630	09	17 40 299	0 052	+62	28 38 45	0 54
DJ 0723-00	07	23	17.854	0 039	-00	48 55 45	0 74	P 0920-39	09	20 48.221	0 046	-39	46 42 35	0 58
OI 446	07	27	24 120	0 045	+40	56 11 0	0 70	P 0922+005	09	22 35 978	0.996	+00	33 25 50	15 88
P 0727-11	07	27	57 994	0 056	-11	34 51 20	0 83	4C 39 25	09	23 55 319	0.036	+39	15 22 43	0 56
GC 0729+25	07	29	52 690	0 041	+25	55 07 84	0 85	MC 0933+119	09	38 31.740	0 044	+11	59 12 51	0 75
GC 0730+50	07	30	04 377	0 052	+50	28 39 88	0 79	VRO 40 09 02	09	45 50 106	0 036	+40	53 43 09	0 54
GC 0731+47	07	31	20.702	0 043	+47	56 44 47	0 64	GC 0954+65	09	54 57 840	0 049	+65	48 15 43	0 59
GC 0733+30	07	33	04 637	0 044	+30	01 04 16	0 87	OK 492	09	55 08.534	0 044	+47	39 28 50	0 60
P 0733-17	07	33	31 244	0 100	-17	29 04 04	1 32	GC 1007+41	10	07 26 129	0 040	-41	47 25 76	0 61
P 0735+17	07	35	14 109	0 027	+17	49 08 88	0 53	OL 318	10	10 54.780	0 044	+35	00 44.05	0 74
P 0736-06	07	36	30 325	0 492	-06	20 05 21	8 08	GC 1013+20	10	13 59 370	0.034	+20	52 47 38	0 51
B2 0738+27	07	38	20 932	0 041	+27	13 48 32	0 56	OL 331	10	18 24 120	0 044	+34	52 28 90	0 75
B2 0742+31	07	42	30 731	0.039	+31	50 16.25	0 64	GC 1019+42	10	19 13.452	0 047	+42	54 34 61	0 72
P 0743-006	07	43	21 055	0 044	-00	36 56 26	0 85	OL 333	10	19 39 881	0 041	+30	56 15 08	0 72
B2 0745+24	07	45	35 789	0 070	+24	07 54.48	1 04	P 1020+191	10	20 11 792	0 041	+19	08 45 48	0 58
OI 478	07	46	39 932	0 050	+48	22 30 51	0 69	GC 1020+40	10	20 14 553	0 041	+40	03 26 83	0 55
P 0748+126	07	48	05 016	0 035	+12	38 45.56	0 94	P 1021-00	10	21 56 201	0 092	-00	37 41.67	1 90
GC 0748+33	07	48	41 037	0 040	+33	21 03 39	0 60	GC 1030+41	10	30 07 796	0 042	+41	31 34 52	0 58
OI 582	07	49	06 446	0 040	+54	00 46.23	0 59	GC 1030+39	10	30 27.529	0 040	+39	51 19 80	0 63
OI-187	07	52	05 439	0.420	-11	39 30 14	7 03	OL 553	10	31 55.974	0 042	+56	44 18 38	0 61
P 0754+100	07	54	22 601	0 043	+10	04 39 40	0 64	P 1034-293	10	34 55.794	0.039	-29	18 26 78	0 54
GC 0802+21	08	02	42 673	0 076	+21	15 27 26	1 14	OL 064 5	10	38 40 908	0 048	+06	25 58 65	0 78
GC 0804+49	08	04	58 388	0 043	+49	59 23.05	0 52	OL 564	10	38 43 161	0.043	+52	49 09 64	0 59
GC 0805+41	08	05	33 657	0 038	+41	01 33 14	0 65	3C 245	10	40 05.956	0 066	+12	19 15 26	0 96
P 0805-07	08	05	49 634	0 050	-07	42 24 02	0 91	OL 569	10	41 07.014	0 046	+53	38 05 63	0 61
P 0808+019	08	08	51 155	0.026	+01	55 50 80	0 73	P 1054+004	10	54 41.940	0 058	+00	28 06 60	1 15

Table 5 (contd)

SOURCE NAME	HR	RIGHT MIN	ASCENSION SEC	ERROR	DEG	MIN	SEC	ERROR	SOURCE NAME	HR	RIGHT MIN	ASCENSION SEC	ERROR	DEG	MIN	SEC	ERROR		
P 1055+20	10	55	37.509	0.039	+20	07	54	73	0.80	P 1302-102	13	02	55.453	0.163	-10	17	10	11	2.38
P 1055+01	10	55	55.320	0.040	+01	50	00	44	3.45	GC 1307+56	13	07	05.094	0.042	+56	13	36.15	0.59	
GC 1058+39	10	58	42.210	0.044	+39	20	39	97	0.70	B2 1308+32	13	08	07.637	0.068	+32	36	38	29	1.01
B2 1101+38	11	01	40.532	0.034	+38	28	42	25	0.51	OP 114	13	08	39.190	0.054	+14	33	42	76	1.53
GC 1104+16	11	04	36.631	0.045	+16	44	16.62	0.72	OP-322	13	13	20.055	0.069	-33	23	09	34	1.02	
GC 1106+38	11	06	43.450	0.045	+38	00	47	12	0.74	OP 326	13	15	17.798	0.031	+34	41	02	63	0.43
OM 316 3	11	09	49.557	0.048	+35	19	58.42	0.86	CC 1325+43	13	25	10.559	0.040	+43	41	59	45	0.55	
GC 1111+14	11	11	21.293	0.038	+14	58	47	97	0.60	P 1327-311	13	27	29.954	0.045	-31	07	31	28	0.56
P 1123+26	11	23	14.839	0.035	+26	26	49	46	0.56	3C 287 1	13	30	20.519	0.060	+02	16	07	61	1.48
GC 1128+38	11	28	12.509	0.041	+38	31	51	97	0.57	OP 151	13	31	10.012	0.043	+17	04	25	40	0.72
P 1128-047	11	28	57.464	0.095	-04	43	45	43	1.68	GC 1333+45	13	33	15.705	0.045	+45	57	56	21	0.66
P 1130+009	11	30	46.169	0.104	+00	57	27	12	1.82	GC 1333+58	13	33	36.446	0.161	+58	59	17	45	1.83
P 1136-13	11	36	38.509	0.078	-13	34	05	40	1.12	P 1333-33	13	33	47.177	0.069	-33	42	39	75	0.89
P 1143-245	11	43	36.362	0.040	-24	30	52.73	0.55	DW 1335-12	13	34	59.772	0.056	-12	42	09	15	0.66	
GC 1144+54	11	44	04.564	0.048	+54	13	22	96	0.59	CC 1335+55	13	35	55.220	0.034	+55	16	15	61	0.45
GC 1144+40	11	44	21.024	0.030	+40	15	14	16	0.47	1336-237	13	36	15.717	0.045	-23	46	01	08	0.56
P 1144-379	11	44	30.956	0.070	-37	55	30.77	1.02	P 1336-260	13	36	32.449	0.049	-26	05	18	18	0.61	
OM-076	11	45	18.141	0.304	-07	07	58	16	4.49	P 1337-033	13	37	37.868	0.051	-03	20	10	94	1.04
P 1146-037	11	46	22.431	0.040	-03	47	30.39	0.81	GC 1337+63	13	37	46.107	0.051	+63	44	08	74	0.60	
1146-037	11	46	23.515	6.465	-03	47	46	60	99.16	GC 1342+662	13	42	17.899	0.057	+66	17	27	75	0.54
P 1148-00	11	48	10.217	0.199	-00	07	41	83	49.55	GC 1342+663	13	42	41.034	0.058	+66	21	12	94	0.54
P 1148-171	11	48	30.248	0.127	-17	07	17	58	1.68	P 1345+12	13	45	06.366	0.162	+12	32	17	09	1.99
OM 484	11	50	47.975	0.041	+49	47	50	08	0.58	GC 1347+53	13	47	42.576	0.030	+53	56	08	34	0.42
GC 1155+48	11	55	52.098	0.042	+48	41	57	62	0.60	P 1348-289	13	48	55.863	0.048	-28	57	29	58	0.62
P 1156-094	11	56	38.936	0.188	-09	24	08	03	2.27	DW 1349-14	13	49	10.786	0.050	-14	34	27	14	0.75
GC 1156+29	11	56	57.764	0.034	+29	31	25	74	0.53	3C 292	13	49	18.777	0.328	+64	46	19	06	2.40
P 1157-215	11	57	18.317	0.039	-21	32	11	51	0.56	P 1349-439	13	49	52.541	0.039	-43	57	54	26	0.55
ON-001	12	00	00.434	0.744	-05	11	19	78	13.36	P 1351-018	13	51	32.120	0.100	-01	51	21	28	1.62
GC 1204+28	12	04	55.092	0.055	+28	11	40	67	0.75	P 1352-104	13	52	06.895	0.188	-10	26	21	87	3.50
B2 1211+33	12	11	32.812	0.047	+33	26	26	11	0.63	P 1354-174	13	54	21.999	0.041	-17	29	24	20	0.56
P 1213-17	12	13	11.692	0.043	-17	15	05	40	0.58	OP-192	13	54	28.564	0.050	-15	12	51	40	0.61
GC 1213+35	12	13	24.812	0.040	+35	04	55	25	0.61	GC 1355+44	13	55	38.215	0.051	+44	08	34	12	0.61
1216-010	12	16	00.977	0.563	-01	03	14	22	9.13	P 1402-012	14	02	11.401	0.063	-01	16	03	48	1.04
ON 428	12	16	38.577	0.039	+48	46	35	23	0.55	P 1402+044	14	02	29.968	0.047	+04	29	55	40	0.77
P 1217+02	12	17	38.302	0.393	+02	20	21	52	7.59	DW 1403-08	14	03	21.605	0.145	-08	33	48	57	2.33
3C 270 1	12	18	03.909	0.017	+33	59	50	53	0.23	P 1404-342	14	04	57.172	0.051	-34	17	14	81	0.61
P 1218-02	12	18	49.941	0.041	-02	25	11	50	0.97	P 1406-076	14	06	17.847	0.097	-07	38	15	16	1.39
P 1219+04	12	19	49.404	0.142	+04	29	51	11	2.76	GC 1406+56	14	06	32.107	0.041	+56	27	43	40	0.56
P 1222+21	12	22	23.405	0.035	+21	39	22	64	0.59	P 1413+135	14	13	33.861	0.065	+13	34	17	87	1.02
3C 272 1	12	22	31.597	0.023	+13	09	49	86	0.43	OQ 323	14	13	56.281	0.070	+34	58	29	47	0.72
ON 343	12	25	30.760	0.036	+36	51	47	11	0.55	4C 46 29	14	15	13.437	0.057	+46	20	55	49	0.79
3C 273	12	26	33.190	0.106	+02	19	44	51	1.72	3C 298	14	16	38.780	0.067	+06	42	20	84	1.02
3C 274	12	28	17.575	0.017	+12	40	01	80	0.26	GC 1418+54	14	18	06.180	0.032	+54	36	58	05	0.43
P 1229-02	12	29	26.299	3.249	-02	07	37	63	51.70	P 1424-41	14	24	46.725	0.071	-41	52	54	44	1.03
P 1236+077	12	36	52.292	0.042	+07	46	45	30	0.68	P 1427+109	14	27	43.692	0.036	+10	56	46	75	0.76
P 1237-10	12	37	05.638	0.301	-10	06	34	94	4.45	OQ 546	14	27	44.043	0.039	+54	19	29	67	0.54
M 104	12	37	24.051	0.074	-11	21	05	05	0.70	OQ-151	14	30	10.654	0.037	-17	48	24	12	0.55
GC 1240+38	12	40	27.031	0.042	+38	07	25	0	0.68	P 1430-155	14	30	36.138	0.058	-15	35	35	10	1.25
ON-073	12	43	28.787	0.100	-07	14	23	37	1.44	P 1435-218	14	35	18.688	0.078	-21	51	57	91	0.81
P 1244-255	12	44	06.720	0.041	-25	31	26	91	0.57	GC 1435+63	14	35	37.228	0.034	+63	49	35	78	0.42
P 1252+11	12	52	07.699	0.063	+11	57	21	47	0.92	P 1437-153	14	37	11.426	0.338	-15	19	00	08	4.55
ON-392	12	55	15.163	0.081	-31	39	04	64	1.10	OQ 363	14	38	22.519	0.058	+38	33	03	16	0.67
ON 393	12	55	35.463	0.049	+32	45	23	0	0.62	OQ 366	14	39	54.033	0.065	+32	47	05	41	0.76
P 1256-220	12	56	13.849	0.044	-22	03	21	22	0.62	3C 303 EAST	14	41	24.831	0.062	+52	14	18	61	0.84
P 1257+145	12	57	51.551	0.041	+14	33	28	59	1.31	P 1441+25	14	41	43.568	0.052	+25	14	24	33	0.62
GC 1300+58	13	00	47.122	0.048	+58	04	43	23	0.61	P 1446-16	14	45	28.334	0.038	-16	07	56	43	0.59
P 1302-035	13	02	08.755	0.050	-03	29	59	06	1.09	P 1449-012	14	49	12.561	0.049	-01	15	18	20	1.16

Table 5 (contd)

SOURCE NAME	HR	MIN	SEC	ASCENSION ERROR	DEG	MIN	SEC	DECLINATION ERROR	SOURCE NAME	HR	MIN	SEC	ASCENSION ERROR	DEG	MIN	SEC	DECLINATION ERROR
P 1451-375	14	51	18	325	0	070			B2 1721+34	17	21	32	026	0	037		
GC 1452+30	14	52	25	043	0	039			GC 1722+40	17	22	27	154	0	057		
MA 1456+04	14	56	29	176	0	030			GC 1725+12	17	25	47	639	0	042		
3C 309.1	14	58	56	591	0	173			P 1725+044	17	25	56	345	0	126		
GC 1459+48	14	59	07	246	0	038			GC 1726+45	17	26	01	190	0	043		
P 1502+036	15	02	35	485	0	223			GC 1732+09	17	32	35	668	0	040		
OR 306	15	04	12	951	0	038			GC 1734+50	17	34	36	730	0	047		
P 1504-167	15	04	16	384	0	068			GC 1739+52	17	39	29	005	0	038		
P 1508-05	15	08	14	986	0	040			GC 1743+17	17	43	22	261	0	036		
P 1511-100	15	11	02	241	0	043			1749+701	17	49	03	426	0	070		
P 1511-210	15	11	03	939	0	040			GC 1751+28	17	51	45	376	0	038		
P 1514+00	15	14	06	816	0	037			OT 486.4	17	51	53	691	0	039		
P 1518+047	15	18	44	877	0	122			P 1756+237	17	56	55	940	0	031		
P 1519-273	15	19	37	286	0	071			OT 398	17	58	44	702	0	039		
OR 342	15	25	16	899	0	079			OU 401	18	00	03	217	0	038		
P 1532+01	15	32	20	250	0	035			P 1806-458	18	06	15	187	0	056		
P 1535+004	15	35	42	576	0	040			GC 1806+45	18	06	56	515	0	039		
GC 1538+14	15	38	30	062	0	075			GC 1807+27	18	07	13	637	0	048		
DW 1543+00	15	43	36	212	0	052			P 1821+10	18	21	41	640	0	048		
P 1546+027	15	46	58	339	0	054			GC 1823+56	18	23	14	950	0	055		
OR 580	15	47	52	270	0	046			GC 1830+28	18	30	52	381	0	042		
DW 1543+05	15	48	06	945	0	058			GC 1842+68	18	42	43	374	0	059		
OR 131	15	48	21	179	0	027			3C 390.3	18	45	37	651	0	151		
P 1550-269	15	50	59	814	0	041			GC 1848+28	18	48	29	082	0	046		
OR 186	15	51	12	021	0	045			3C 395	19	01	02	320	0	037		
P 1555-140	15	55	33	722	0	052			OV-213	19	08	12	599	0	060		
P 1556-245	15	56	41	217	0	063			OV-214	19	08	55	657	0	175		
B2 1600+33	16	00	11	905	0	041			OV-236	19	21	42	437	0	211		
GC 1604+31	16	04	10	594	0	041			P 1922-341	19	22	00	987	0	043		
P 1604-333	16	04	22	159	0	042			OV 239.7	19	23	49	766	0	035		
P 1606+10	16	06	23	384	0	037			GC 1926+61	19	26	49	649	0	050		
P 1614+051	16	14	09	088	0	042			P 1936-15	19	36	36	118	0	052		
4C 36.27	16	15	06	135	0	043			P 1937-101	19	37	12	665	0	064		
P 1615+029	16	15	19	141	0	045			OV 080	19	47	40	528	0	312		
D'J 1616+06	16	16	36	581	0	034			P 1953-325	19	53	48	363	0	235		
P 1622-253	16	22	44	123	0	095			OV 591	19	54	22	488	0	040		
P 1622-29	16	22	57	241	0	045			P 1954-388	19	54	39	009	0	043		
DA 411	16	24	18	271	0	042			P 2000-330	20	00	13	015	0	038		
P 1625-141	16	25	56	548	0	425			P 2008-159	20	08	25	909	0	092		
GC 1633+38	16	33	30	625	0	041			OW-015	20	08	33	727	0	033		
GC 1635+47	16	36	19	159	0	044			P 2012-017	20	12	39	820	0	135		
GC 1637+57	16	37	17	409	0	046			OW 637	20	21	13	380	0	058		
3C 345	16	41	17	630	0	040			P 2029+121	20	29	32	626	0	071		
3C 346	16	41	34	434	0	030			OW 551	20	30	29	133	0	033		
P 1647-296	16	47	29	449	0	047			OW 154.9	20	32	58	529	0	059		
P 1648+015	16	48	31	563	0	045			P 2037-253	20	37	10	747	0	040		
DA 426	16	52	11	739	0	026			OW-174	20	44	30	850	0	037		
DW 1656+05	16	56	05	593	0	042			GC 2047+09	20	47	20	777	0	066		
CC 1656+34	16	56	12	266	0	054			P 2047+039	20	47	36	001	0	096		
1656+48	16	56	24	901	0	039			P 2053-044	20	53	12	821	0	029		
1656+47	16	56	43	969	0	048			P 2053-323	20	53	20	802	0	056		
P 1657-261	16	57	47	698	0	044			P 2058-297	20	58	00	910	0	041		
P 1705+018	17	05	02	784	0	043			P 2059+034	20	59	08	126	0	051		
OT-111	17	06	40	114	0	041			P 2106-413	21	06	19	396	0	042		
GC 1714+21	17	14	03	727	0	058			OX 131	21	18	41	548	0	041		
GC 1716+68	17	16	27	782	0	070			P 2123-463	21	23	13	267	0	038		
GC 1717+17	17	17	00	339	0	034			R 2126-15	21	26	26	759	0	043		

Table 5 (contd)

SOURCE NAME	HR	MIN	SEC	RIGHT ASCENSION ERROR	DEG	MIN	SEC	DECLINATION ERROR	SOURCE NAME	HR	MIN	SEC	RIGHT ASCENSION ERROR	DEG	MIN	SEC	DECLINATION ERROR							
P 2126-185	21	26	33	890	0	043	-18	34	32	51	0	60	P 2312-319	23	12	06	365	0.028	-31	55	01	31	0	39
P 2127-096	21	27	38	415	0	053	-09	40	50	29	0	90	GC 2318+04	23	18	12	148	0.030	+04	57	22	86	0	65
P 2127+04	21	28	02	673	0	047	+04	49	03	35	0	92	P 2318-195	23	18	14	337	0.177	-19	35	58	10	2	32
P 2128-12	21	28	52	698	0	046	-12	20	20	76	0	72	GC 2319+31	23	19	28	291	0.055	+31	47	39	98	0	67
P 2131-021	21	31	35	227	0	126	-02	06	41	57	2	04	B2 2319+27	23	19	32	005	0.039	+27	16	18	66	0	58
P 2135-248	21	35	45	384	0	040	-24	53	28	49	0	56	OZ 533	23	20	06	010	0.045	+50	41	24	03	0	64
P 2137+20	21	37	27	860	0	057	+20	57	57	89	1	16	P 2320-021	23	20	30	352	0.192	-02	07	15	02	3	61
P 2140-048	21	39	59	970	0	078	-04	51	27	59	1	18	P 2320-035	23	20	57	380	0.052	-03	33	31	93	0	90
OX-173	21	43	38	849	0	052	-15	39	37	11	0	71	P 2325-150	23	25	11	594	0.045	-15	04	27	26	0	60
OX 074	21	44	42	487	0	066	+09	15	50	91	1	00	GC 2327+33	23	27	45	965	0.037	+33	32	03	46	0	55
P 2145+06	21	45	36	566	0	086	+06	43	33	18	1	39	P 2328+10	23	28	08	805	0.055	+10	43	45	16	0	78
P 2147+14	21	47	59	332	0	095	+14	35	44	85	1	14	OZ 347	23	28	17	891	0.051	+31	39	00	25	0	71
P 2149-306	21	49	00	619	0	039	-30	42	00	80	0	55	P 2329-16	23	29	02	119	0.082	-16	13	26	82	1	16
OX 081	21	49	02	057	0	057	+06	55	21	11	0	86	P 2329-384	23	29	18	905	0.038	-38	23	21	47	0	53
OX 082	21	49	07	663	0	047	+05	38	07	38	0	71	P 2330+083	23	30	25	058	0.033	+08	21	35	74	0	58
P 2149+17	21	50	02	236	0	036	+17	20	29	58	0	55	OZ-252	23	31	17	903	0.044	-24	00	15	40	0	57
OX-192	21	55	22	322	0	294	-15	15	15	63	4	59	P 2335-18	23	35	20	659	0.028	-18	08	57	66	0	46
P 2200-238	22	00	07	743	0	037	-23	49	41	67	0	56	P 2335-027	23	35	23	076	0.047	-02	47	32	53	0	98
VRO 42 22 01	22	00	39	387	0	038	+42	02	08	57	0	55	3C 465	23	35	59	006	0.041	+26	45	16	21	0	56
GC 2201+17	22	01	03	350	0	028	+17	11	15	38	0	45	P 2337-334	23	37	16	646	0.040	-33	26	54	66	0	53
P 2201+04	22	01	46	297	0	267	+04	25	30	31	4	58	GC 2337+26	23	37	58	271	0.042	+26	25	18	71	0	72
OY 313	22	07	34	085	0	036	+35	41	14	93	0	53	GC 2338+33	23	38	27	403	0.040	+33	02	24	44	0	56
P 2208-137	22	08	42	892	0	037	-13	42	59	46	0	60	P 2344+09	23	44	03	773	0.032	+09	14	05	51	0	46
P 2209+08	22	09	32	269	0	039	+08	04	25	44	0	78	HIA 2344+092	23	44	32	986	0.076	+09	14	05	48	1	17
P 2209+236	22	09	45	661	0	053	+23	40	50	35	0	60	P 2349-01	23	49	22	357	0.085	-01	25	55	43	1	84
P 2210+01	22	10	05	067	0	238	+01	38	00	92	4	01	OZ 486	23	51	49	977	0.042	+45	36	22	79	0	55
GC 2214+35	22	14	07	035	0	028	+35	03	15	01	0	39	OZ-187	23	51	55	724	0.099	-15	29	50	87	1	42
P 2215+02	22	15	15	707	0	427	+02	05	07	13	7	11	P 2352-04	23	52	17	807	0.043	-04	21	45	32	1	01
P 2216-03	22	16	16	406	0	042	-03	50	41	07	0	74	DA 611	23	52	37	795	0.050	+49	33	26	60	0	73
P 2221-116	22	21	28	616	0	052	-11	41	35	70	1	03	P 2354-11	23	54	55	762	1.555	-11	41	58	52	24	54
3C 446	22	23	11	045	0	043	-05	12	17	43	0	74	P 2355-106	23	55	36	969	0.042	-10	36	50	89	0	67
P 2223+21	22	23	14	768	0	054	+21	02	50	03	0	83	P 2356+196	23	56	12	584	0.038	+19	38	38	36	0	67
P 2227-08	22	27	02	263	0	053	-08	48	16	47	0	83	GC 2356+38	23	56	59	832	0.072	+38	34	00	23	0	65
P 2227-399	22	27	44	970	0	105	-39	58	16	23	1	15												
P 2229-17	22	29	41	003	0	081	-17	14	29	78	1	10												
P 2233-148	22	33	53	999	0	026	-14	48	57	18	0	42												
GC 2234+28	22	34	01	726	0	036	+28	13	23	20	0	51												
GC 2236+12	22	36	06	142	0	085	+12	27	12	73	1	16												
P 2239+096	22	39	19	853	0	031	+09	38	09	87	0	54												
OY-268	22	40	41	708	0	040	-26	00	14	17	0	56												
P 2245-128	22	45	14	125	0	035	-12	53	11	37	0	59												
P 2245+029	22	45	26	015	0	048	+02	54	51	08	1	69												
GC 2246+20	22	46	34	629	0	069	+20	51	11	66	0	83												
P 2247+13	22	47	16	051	0	035	+13	15	15	84	0	56												
GC 2251+24	22	51	44	393	0	034	+24	29	24	16	0	55												
GC 2251+13	22	51	51	879	0	055	+13	25	48	73	0	88												
P 2252-089	22	52	27	533	0	027	-09	00	05	40	0	45												
GC 2253+41	22	53	19	857	0	042	+41	46	51	64	0	63												
P 2254+024	22	54	44	626	0	037	+02	27	13	85	1	03												
OY 492	22	55	04	662	0	040	+41	38	13	21	0	49												
P 2255-282	22	55	22	460	0	037	-28	14	25	72	0	55												
2256+017	22	56	24	708	0	229	+01	47	35	38	3	92												
P 2300-307	23	00	22	087	0	039	-30	46	21	59	0	56												
P 2300-18	23	00	23	374	0	155	-18	57	34	42	1	75												
P 2303-052	23	03	40	121	0	077	-05	16	02	02	1	58												
P 2307+10	23	07	57	555	0	030	+10	39	12	86	0	53												
P 2311-452	23	11	21	897	0	030	-45	12	10	78	0	40												

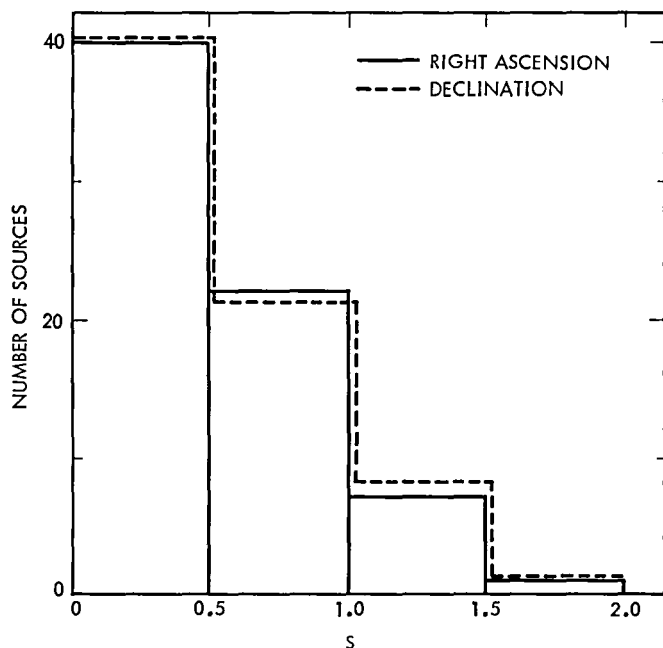


Fig 1 Comparison of position estimates for 70 multiply observed sources. For each source, S is the ratio of the weighted rms scatter of individual position coordinate estimates about the mean value to the weighted rms estimate of the uncertainty in that position coordinate

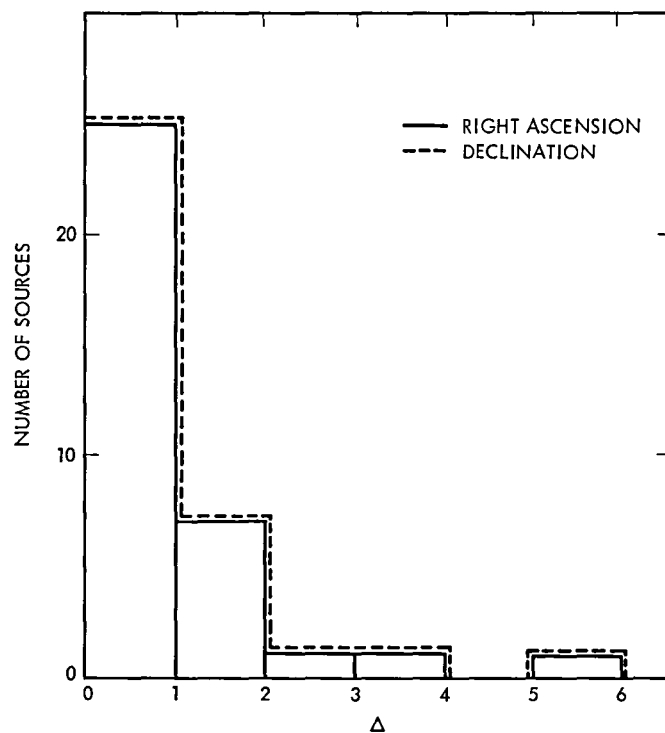


Fig 2 Comparison of position estimates with more accurate radio positional catalogs. For each source, Δ is the ratio of the absolute value of the difference between our source coordinate estimate and a more accurate value to our estimate of uncertainty in that coordinate

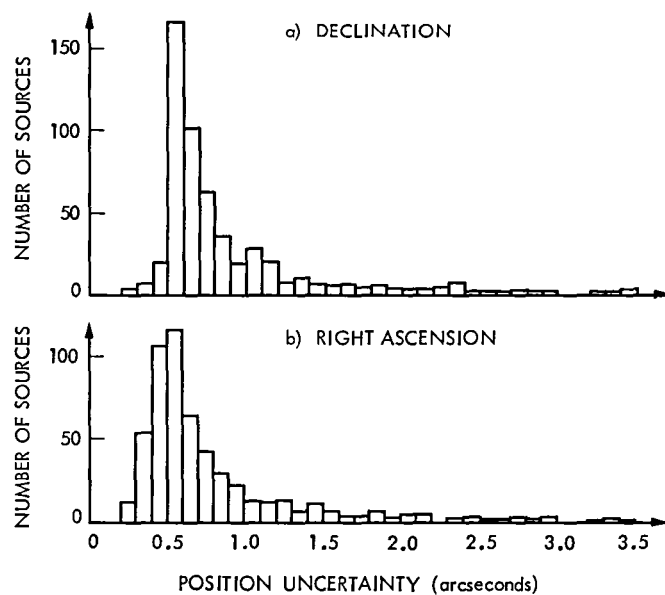


Fig. 3. Histogram of the number of sources versus estimated position uncertainty

An FM/CW Method for the Measurements of Time Delays of Large Cassegrain Antennas

T Y Otoshi

Radio Frequency and Microwave Subsystems Section

An FM/CW measurement technique, which utilizes the time domain principle, was developed to measure the time delays of large Cassegrain antennas. Measured primary path delays of 60 ns (or greater) typically agreed with calculated delays to within 3 ns. Multipath sources generated within the antenna optics region have been identified.

I. Introduction

Large Cassegrain microwave antennas are currently being used as the ground station terminal in most deep space communication systems. At some stations the same antenna is also used as part of a ranging system to obtain spacecraft-to-Earth range data needed for spacecraft navigation. To determine the true range of the spacecraft from the ground station location, time delay calibrations must be performed on the microwave antenna as well as on the ground station equipment (Ref 1). Accurate determination of the signal delay time through the antenna is also critical to the success of clock synchronization experiments using Very Long Baseline Interferometry (VLBI) techniques (Ref 2).

Extensive experimental work has been done in the JPL/NASA Deep Space Network (DSN) to develop a simple but accurate method for the measurement of antenna time delay of a large Cassegrain antenna. Most of the methods investigated were found to be susceptible to large errors caused by multipath signals generated within the region between the main reflector and subreflector surfaces. This paper describes

a time domain measurement technique which enables accurate measurement of the primary signal delay time without interference from major multipath signals. The following sections of this paper present discussions of the multipath problems associated with antenna time delay calibrations, FM/CW theory and instrumentation, and experimental data obtained on 64- and 26-m antennas.

II. Antenna Delay Calibration Problems

Figure 1 shows a typical 64-m-diameter Cassegrain antenna used in the DSN. This particular antenna is equipped with a reflex-dichroic system to enable simultaneous reception of S- and X-band frequencies. The feedcones are aluminum housings that contain feed assemblies as well as transmitter and receiver equipment. A signal originating from the far-field is collected by the parabolic reflector, reflected to the subreflector, re-reflected, and arrives at the receive horn via the ray paths of the dual-band system.

The calibration problem is to determine the antenna delay, which for purposes of this paper shall be defined as the time it takes for the illuminating signal to propagate from a designated reference plane (such as the antenna focal plane) to a designated output port of the receive horn assembly. It should be pointed out that another definition of antenna delay, which does not include the waveguide delays of the feed horn assembly, was given previously in Ref. 3.

A method that has been used for the calibration of antenna delay (as defined in this paper) is depicted in Fig. 2. The horn on the dish surface can be viewed as one "cell or element" of the total signal collected by the many cells of the parabolic reflector. The calibration procedure is to inject a signal into the dish-mounted horn that is pointed toward the subreflector. A time delay measurement is made between the input port of the transmit horn (on dish surface) and a designated output port in the receive system horn assembly. A theoretical constant is then added to the measured value to artificially move the signal input reference plane from the dish location to the paraboloid focal plane. In the ideal case, the corrected measured delay value would be the same as the antenna delay for the entire Cassegrain antenna. The basic assumption is made that all signals follow geometric optics (GO) ray paths so that they all have the same delay (path length) from the paraboloid focal plane to the system feed horn focal point.

In practice numerous factors such as blockage, mechanical supports, imperfect reflector surfaces, mechanical distortions, and defocusing cause departure from the ideal Cassegrain antenna. Furthermore, two known types of multipath signals are generated within the antenna optics media of practical Cassegrain antennas. The first type is caused by a portion of the illuminating plane wave signal becoming blocked by the large subreflector and subreflector support struts. Diffraction around the edges of the subreflector and struts will occur and a portion of these diffracted fields propagates directly into the receive horn. Another type of multipath phenomenon that occurs is that generated by unwanted multiple reflections between the subreflector and feedcone surfaces (Fig. 3).

In order to determine the magnitude of deviations from ideal antenna time delay caused by the multipath phenomenon, far-field time delay experiments were performed on DSN 64-m antennas. The far-field signals were ranging signals provided by Viking spacecraft at S-band and Voyager spacecraft at X-band. From the test data (Ref. 4) it was found that multiple-reflection type multipath signals produced maximum errors of ± 3 ns when a ranging modulation frequency of 500 kHz was used. These errors were produced by multipath signals at least 40 to 46 dB down from the primary signal for the 64-m antenna. These experimental observations were also supported by results of a far-field theoretical study performed

by Rusch (Ref. 5). In a separate study, a cursory theoretical analysis showed that diffractions from the subreflector edge and struts would produce time delay errors smaller than those produced by multiple reflections. Although these far-field experiments and analyses provided valuable information on the effects of multipath errors on the total Cassegrain antenna received signal, there was still a need for an accurate method to calibrate *absolute* antenna delay.

During the period of 1975-1976, several methods were investigated as possible ways to calibrate absolute antenna and ground station delays. Far-field methods for calibrating absolute antenna delays were shown by Sato in Ref. 6 to be expensive, and it was concluded that the necessary technology and equipment would take considerable effort and time to develop. As a parallel effort to find different methods to calibrate antenna delays, a separate theoretical study was made by A. Cha (Ref. 7) to investigate near-field methods as well. The results from this theoretical study showed that if multipath effects were negligible, then far-field absolute antenna delays for 26-m and 64-m DSN-type antennas could be determined to a 1 or 2 ns accuracy level with the horn-on-dish calibration method. He found that for propagation paths between the main reflector surface and the system feed horn, the delays computed by the geometric theory of diffraction (GTD) agreed with the delays computed by the geometric optics theory to better than 1 ns. It was concluded from this study that further efforts should be directed towards developing the horn-on-dish method for absolute antenna delay calibrations.

A major problem discovered after extensive testing with the horn-on-dish method was the generation of new multipath signals by the calibration geometry itself. Very often the antenna pattern of the transmit horn (on dish surface) will have insufficient directivity and high side-lobes. Therefore, the calibration signal will not only follow the primary ray path, but will also follow other ray paths due to unintentional illuminations of the struts and side of the feedcone housings. When conventional time delay measurement methods (modulation and phase-slope methods) were employed to measure time delays, it was found that the accuracy of results was seriously degraded by multipath signals and critically dependent upon (1) the directivity of the calibration horn used, (2) the selected location of this horn on the dish surface, and (3) the bandwidth of the measurement.

After experiencing considerable problems with conventional techniques, a solution was made possible by a diagnostic instrument that utilized the FM/CW technique based on the time domain principle. Not only did it become possible to isolate and make accurate measurement of primary path time delays, but it also became possible to positively identify the

existence and origins of multipath signals associated with any calibration configuration

III. FM/CW Theory and Instrumentation

The instrument used for making antenna time delay measurements is a Scientific-Atlanta Series 1691 Fault Locator (Ref 8), which is an FM/CW radar system (Ref 9) that can be operated at any center frequency in the microwave range from 1.7 to 12.4 GHz with sweep bandwidths varying from 40 to 100 MHz. The instrument is intended primarily to be used as a diagnostic tool to locate various discontinuities in a microwave transmission line and measure the magnitude of return losses as functions of the distances (range) to the discontinuities. For the test results of this paper, however, the instrument was modified to operate in a "transmission" or one-way mode rather than in the normal return loss or bistatic radar mode. Figure 4 shows the instrument with the modifications. In this transmission mode, the measured delays can be compared to path lengths obtained from antenna mechanical drawings, and the measured magnitudes of the received signals can be compared to those calculable from the Friis transmission formula (Ref 10).

Figure 5 is a functional block diagram of the instrumentation. A sample of the oscillator output signal is fed to the local oscillator input of the mixer. The remainder of the signal is fed to the transmit antenna. Figure 6 shows that the oscillator output frequency is changed in a sawtooth manner, over the sweep bandwidth $2\Delta F$, around the center frequency F_0 . Thus, data is gathered over the $F_0 \pm \Delta F$ frequency range. The solid line represents the oscillator output (transmit signal) as a function of time while the dashed line represents the received signal. Only one of the receive signals is shown in Fig 6 in order to illustrate the basic principle of operation. The basic relationship for this instrumentation as derived by the proportional triangle relationships in Fig 6 is

$$t_g = f_d \left(\frac{T}{2\Delta F} \right) \quad (1)$$

where t_g is the propagation time or time delay of interest, f_d is the difference frequency, T is the sawtooth modulation period, and ΔF is one-half the sweep bandwidth. If d is defined as the path length from transmit antenna to the receive antenna, and the group velocity v_g of the signal in the propagation media is known, then the path length can be computed from

$$d = v_g f_d \left(\frac{T}{2\Delta F} \right) \quad (2)$$

If f_d is fixed, then the path length is inversely related to the oscillator modulation rate, $2\Delta F/T$.

The mixer (see Fig 5) produces an output signal that contains a difference frequency f_d for each received signal of different path length. The control unit of this instrument varies the sweep bandwidth and causes the distance readout to indicate a focus distance d when the difference frequency produced by the received signal is equal to 5 kHz and has an integer number of cycles in the sawtooth period T shown in Fig 6. The narrowband IF amplifier rejects all the other received "target" signals and feeds this signal into the log converter which drives the signal magnitude meter. Since the distance indicated by this instrument is calibrated for an equivalent bistatic radar configuration, the "target" distance readings are multiplied by 2 when the instrument is modified for use for the one-way or transmission mode.

It is beyond the scope of this article to present further details of the instrument. The mathematical relationships of the FM/CW theory can be found in Refs 8 and 9 and additional details on the Fault Locator design can be found in Scientific-Atlanta Fault Locator instrument manuals. It is sufficient for purposes of this article to state that when the Fault Locator is focused at a distance in the vicinity of a single "target" response, the signal magnitude meter and plot recorder output displays a response of the form

$$R_{dB} = 20 \log_{10} \left| \frac{A \sin(\pi u)}{\pi u (1 - u^2)} \right| \quad (3)$$

where A is the target signal amplitude normalized to a reference signal level that was initially established with a zero dB set control. In Eq (3), u is a variable that is directly related to sweep bandwidth and to the difference between the focus distance and the "target" distance. A peak response is observed when u becomes zero as a result of adjusting the focus distance to be equal to the "target" distance. The measured propagation path length at this peak response is the distance between the transmit and receive antennas, provided that the instrumentation has been adjusted to indicate zero when there is no separation between the antennas.

Increasing the sweep bandwidth narrows the width of the response and improves the resolution of the delay measurement. However, the sweep bandwidth cannot be controlled in the commercial units due to automatic setting of the sweep bandwidth to a value (between 40 to 100 MHz) depending upon the target range.

Further work has since been done by the Howland Company of Atlanta, Ga., to enable manual setting of sweep band-

widths, increasing sweep bandwidths up to 300 MHz, and increase the range of transmission mode measurements from 1500 to 30,000 ns (Ref 11)

IV. Measurement of Primary Path Delays

Figure 7 shows a block diagram of the test setup that was used to measure time delays of primary and secondary signals within the antenna optics media of a Cassegrain antenna. The modified Fault Locator instrument and recorder were placed inside the S-band Cassegrain feedcone close to the base of the receive system feed horn. A short 1.8-m (6-ft) cable was connected between the receive system output port and the Fault Locator "receive" port. Then a 30.48-m (100-ft) low-loss flexible cable with amplifiers in series was connected between the Fault Locator signal output port and the transmit horn (standard gain type) located on the dish surface as shown in Fig 7. The amplifiers were of the compact solid-state type and provided a combined gain of about 40 dB to compensate for the airpath loss between the transmit horn and the system receive horn. All components including test cables and amplifiers were pretested and selected for constant group delay over a wide bandwidth and also for group and phase delay stability with mechanical flexing and temperature changes.

The basic measurement technique is described as follows. A measurement is first made of the total system delay. The total system for this test configuration consists of the transmit horn, the airpath, receive horn feed assembly, cables, amplifiers, and the Fault Locator instrument. Then a reference measurement is made by disconnecting (1) the cable at the transmit horn and (2) the cable at the receive system output port, and then connecting the two cables (with amplifiers) together with a 40-dB coaxial pad inserted in between. The 40-dB pad value was used to attenuate the increase in signal level due to removal of the airpath loss which was in the total system measurement. Subtraction of the reference measurement from the total system measurement, with small corrections applied for the delay of the transmit horn and 40-dB pad, resulted in the primary path delay value of interest.

To make the actual measurement of primary path delay with the Fault Locator instrument, the dial on the Fault Locator is manually adjusted to a setting that provides maximum received signal. The path length is directly read off the dial in units of feet and multiplied by 2 for transmission mode measurements. For these antenna delay measurements, the resolution of the measurements was improved to about 3.1 cm (0.1 ft) by incorporating more sensitive indicators into the instrument. All measured values were later converted from path length in feet into time delay in ns.

Measurements through the use of the described test procedure were made on the 64-m antenna at S- and X-band frequencies with the transmit horn at the different test locations shown in Fig 8(a). With the exception of Location 5, all other locations were on the dish surface. Location 5 was adjacent to and on the same level as the tricone support platform (see Fig 2) and therefore physically closer to the subreflector than the other locations. Table 1 presents a summary of the measured primary path delays and deviations from calculated values. The primary path delays are absolute time delays for a signal propagating from the horn-on-dish location to the output flange of the receive system horn assembly via the Cassegrain system primary ray paths. The calculated primary path delay values in Table 1 were based on physical dimensions for the airpath and group velocities of various waveguides in the receive horn feed assembly. It can be seen in Table 1 that all of the measured primary path delays were greater than 100 ns and typically agreed with calculated values to within 3 ns. About 1 ns of the disagreement is attributed to uncertainties in the calculated values. The results of the calculated feedhorn assembly delays are tabulated in Ref 12.

Table 2 shows the results of similar measurements of primary path delays performed on a 26-m DSN antenna. The experimental setup was similar to that for the 64-m antenna except that a shorter cable was used to connect the Fault Locator to the transmit horn on the dish surface at the different locations shown in Fig 8(b). As may be seen in Table 2, the measured primary path delays for the 26-m antenna were between 60 and 75 ns. The agreement between measured and calculated delays was typically within 3 ns. The calculated delays were again based on physical dimensions from antenna drawings and group velocities of the waveguide components of the feed assembly.

The purpose of comparing calculated and measured values is to provide a basis for evaluating the accuracy of antenna delay measurements when performed with the FM/CW technique. Previous attempts to measure the same delay paths with other measurement techniques led to disagreement with calculated values as large as 20 to 30 ns in some cases (Ref 13). The order-of-magnitude improvement of accuracy obtained with the FM/CW method is basically due to the time domain principle and large sweep bandwidths of 40 to 100 MHz which enable isolation of the primary signal from the secondary signals during the measurement.

For both the 64- and 26-m antennas, it was fortunate that the feed assembly consisted of relatively simple components whose delay could be calculated. In the DSN, there are more complex Cassegrain feed systems such as ones with a monopulse feed whose delay is difficult to calculate accurately. For such complex feed systems, a relatively simple and direct mea-

surement of antenna delay with the FM/CW instrument can now be performed

V. Identifications of Multipath Sources

Identification of multipath signals is easily accomplished by means of a time domain recording such as the plot shown in Fig 9. With the horn-on-dish calibration setup discussed in the previous section, this recording was made automatically with a push of a button on the Fault Locator instrument. The plot not only displays the primary signal response but the responses of all other secondary signals. The horizontal scale of Fig 9 was corrected so that it shows absolute path length in the transmission mode. The multiple bounce surfaces for multipath sources labeled a, b, and c in Fig 9 are identified in Fig 3. Since multiple reflection signals travel longer distances to arrive at the receive horn, they appear on the plot at path lengths longer than the primary signal. It is of significance to note that the relative magnitudes of the largest multipath signals are approximately 20 dB down from the primary signal. Also note that the relative position of the Cassegrain subreflector has a large effect on the magnitudes of the multipath signals, but very little effect on the magnitude of the primary signal. Movement of the subreflector and observations of the changes of the sidelobe structures provided a means for verifying that the sidelobe responses were due to multipath signals and were not part of the normal $(\sin \pi u)/(\pi u)$ type response of the primary signal (See Eq 3).

Similar types of recordings were obtained at all of the horn locations and microwave frequencies shown tabulated in Table 1. In order to verify that the multipath sources were correctly identified, a 2-ft-square absorbing sheet was systematically placed at multiple reflection points such as at the a, b, c locations shown in Fig 3. If the multiple reflection point of the multipath source was correctly located, a new recording would show a decrease of 1 to 2 dB in the response of the multipath signal being investigated. This simple procedure was used to confirm the location of the multiple reflection surfaces on the antenna structure.

Figure 10 shows similar types of recordings taken on a 26-m DSN antenna. Figure 10(a) is a plot taken with the transmit horn near the side of the feedcone housing. At this horn

location, the strength of the largest multipath signal was about 28 dB down from the peak response of the primary signal. Figure 10(b) is a plot taken with the transmit horn placed near the outer edge of the reflector surface. It is interesting to note that at this location the dominant multipath signal has increased relative to the primary signal.

Recordings taken for various horn locations on the reflector surface proved to be very useful for providing insight into the mechanism of the generation of multipath signals and understanding of multipath phenomena for the far-field signal case as well. The diagnostics work with the Fault Locator confirmed that the horn-on-dish method would be very susceptible to large errors if used with the DSN ranging system, which operates with a small modulation bandwidth of 500 kHz. This work led to the development of a "translator method," which is described in Ref. 1 and is a new method now being used throughout the DSN to calibrate ground station delays for ranging systems.

VI. Conclusions

The FM/CW technique in conjunction with the horn-on-dish configuration proved to be the most rapid, convenient, and accurate method for calibrating absolute time delays for a large antenna. The work discussed in this article was performed in 1977 and the results at that time represented a major improvement in achievable accuracies in the measurements of signal path delays within the antenna optics media of large Cassegrain antennas. Through the use of a commercially available FM/CW instrument, accuracies of about 3 ns were obtained on measurement of primary path delays of 60 ns or greater.

The FM/CW instrumentation was invaluable as a diagnostics tool and facilitated increased understanding of the mechanisms of multipath signal generation within the antenna optics media. Time domain plots enabled positive identifications of major multipath signals that could be deleterious to accuracy of antenna time delay calibrations. A beneficial outcome of this work is that attention is now being focused on minimizing multipath signals on future DSN antennas to be used for ranging and VLBI applications.

Acknowledgment

The author thanks R. B. Lyon and M. Franco of JPL for their assistance in setting up the experiment and collecting the data. A. Ray Howland, president of the Howland Co. in Atlanta, Ga., suggested the use of the Fault Locator instrument to locate multipath sources within the antenna propagation paths. He also consulted on the instrument modification and assisted with the measurements and interpretation of the test data.

References

- 1 Komarek, T., and Otoshu, T. Y., "Terminology of Ranging Measurements and DSS Calibrations," in *The Deep Space Network Progress Report 42-36*, pp. 35-40, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1976.
- 2 Otoshu, T. Y., "Definition of Antenna Microwave Time Delay for VLBI Clock Synchronization," in *The Deep Space Network Progress Report 42-49*, pp. 45-56, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1979.
- 3 Cha, A. G., Rusch, W. V. T., and Otoshu, T. Y., "Microwave Delay Characteristics of Cassegrainian Antennas," *IEEE Trans. Ant. Prop.*, Vol. AP-26, pp. 860-865, Nov. 1978.
- 4 Otoshu, T. Y., and Brunn, D. L., "Multipath Tests on 64-m Antenna Using the Viking Orbiter -1 and -2 Spacecraft as Far-Field Illuminators," in *The Deep Space Network Progress Report 42-31*, pp. 41-49, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1976.
- 5 Otoshu, T. Y., and Rusch, W. V. T., "Multipath Effects on the Time Delays of Microwave Cassegrainian Antennas," *AP-S International Symposium Digest - Antennas and Propagation*, Vol. II, 80 CH 1557-8AP, pp. 457-460, June 1980.
- 6 Sato, T., "Feasibility Study of Far-Field Methods for Calibrating Station Delays. An Interim Report," in *The Deep Space Network Progress Report 42-41*, pp. 51-56, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1977.
- 7 Cha, A. G., private communication, June 1977.
- 8 Howland, A. Ray, "Testing Microwave Transmission Lines," *IEEE S-MTT Conference Digest 1974*, pp. 258-260, June 1974.
- 9 Ismail, M. A-E. W., "A Study of the Double Modulated F. M. Radar," Thesis presented to Swiss Federal Institute of Technology, Zurich, 1955.
- 10 Kraus, J. D., *Antennas*, McGraw-Hill, New York, p. 54, 1950.
- 11 Howland, A. R., "Modified Fault Locator Evaluation," *Howland Company Report HC-685115*, Howland Company, Atlanta, Ga., Nov. 14, 1978.
- 12 Otoshu, T. Y., Wallace, K. B., and Lyon, R. B., "Dual Coupler Configuration at DSS 14 for the Voyager Era," in *The Deep Space Network Progress Report 42-42*, pp. 184-190, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1977.
- 13 Otoshu, T. Y., "A Collection of Articles on S/X-Band Experiment Zero Delay Ranging Tests," *Technical Memorandum 33-747*, Vol. 1, Jet Propulsion Laboratory, Pasadena, Calif., p. 123, Nov. 1975.

Table 1 Results of 64-m antenna primary path delay measurements

Horn-on-dish location	System horn polarization	Test frequency, GHz	Measured delay, ns	Measured minus calculated delay, ns
1	LP	2 113	172 3	0 5
2	LP	2 113	164 7	1 9
2	RCP	2 113	164 1	1 3
3	LP	2 113	159 0	3 5
3	RCP	2 113	157 1	1 6
4	RCP	2 113	166 4	3.7
5	LP	2 113	129 3	-0 8
5	RCP	2 113	132.2	2.0
<hr/>				
1	LP	2 295	169 0	-0 2
1	RCP	2 295	170 1	0.9
3	LP	2 295	154.9	2 1
3	RCP	2 295	154 9	2 1
4	RCP	2 295	160 0	-0 1
5	LP	2 295	129 3	1 7
5	RCP	2 295	126 7	-0 9
<hr/>				
1	RCP	8 415	143 4	0 8
2	RCP	8 415	136 8	3.1
3	RCP	8 415	130 1	3 6
4	RCP	8 415	135 4	1 7
5	RCP	8 415	100 9	-0.3

Table 2. Results of 26-m antenna primary path delay measurements

Horn-on-dish location	System horn polarization	Test frequency, GHz	Measured delay, ns	Measured minus calculated delay, ns
1	LP	2 113	62 4	0 4
1	RCP	2 113	65 4	3 5
2	LP	2 113	72 2	0 4
2	RCP	2 113	72 5	0 7
<hr/>				
1	LP	2 295	63 0	1.3
1	RCP	2 295	62 6	0 9
2	LP	2 295	74 2	2 7
2	RCP	2 295	74 9	3 4
3	LP	2 295	68 1	-1 2

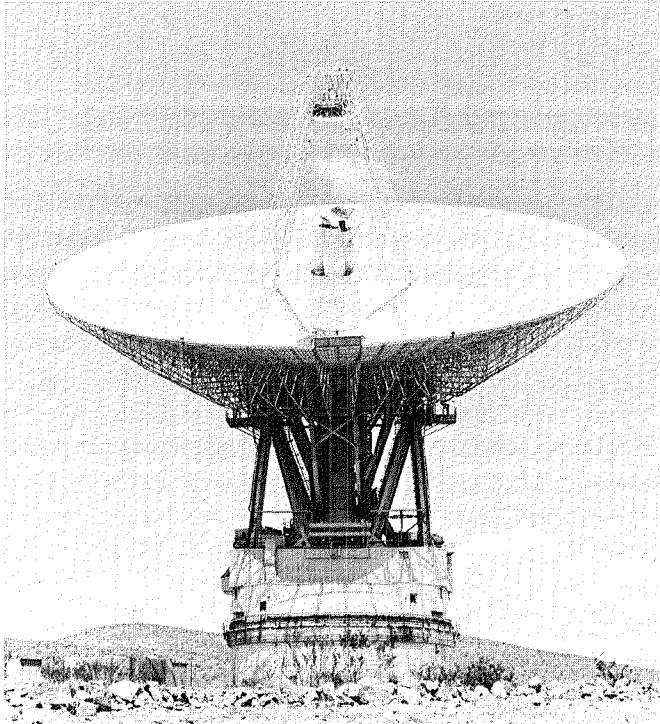


Fig. 1. DSN 64-m Cassegrain antenna at Goldstone, Calif.

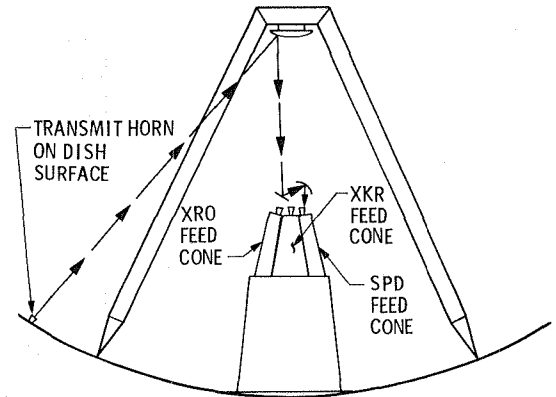


Fig. 2. Principal calibration system propagation path on DSN 64-m antenna

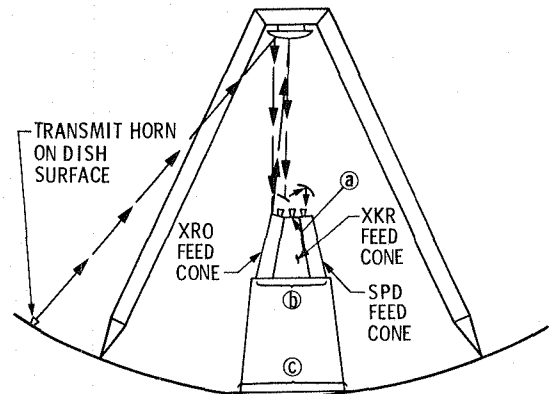


Fig. 3. A secondary propagation path on DSN 64-m antenna

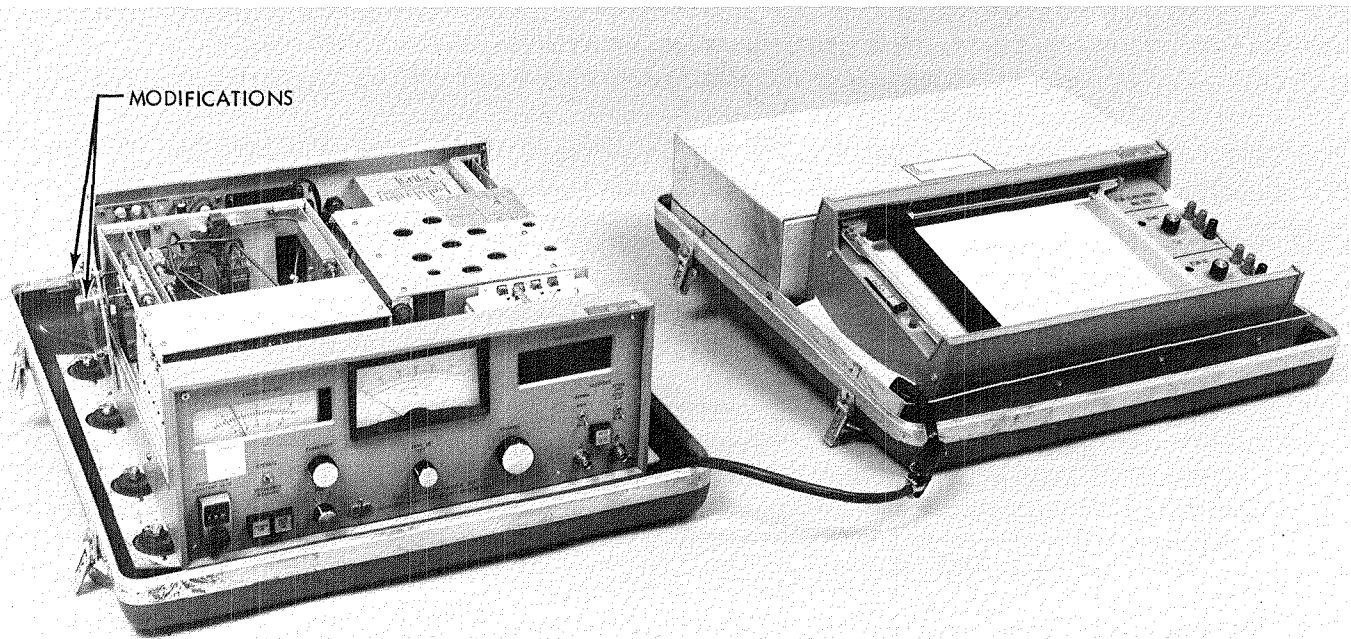


Fig. 4. Modified Fault Locator and recorder

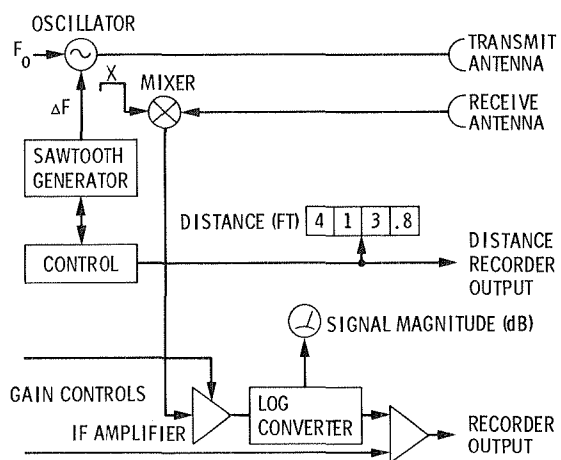


Fig. 5. Block diagram of FM/CW instrumentation

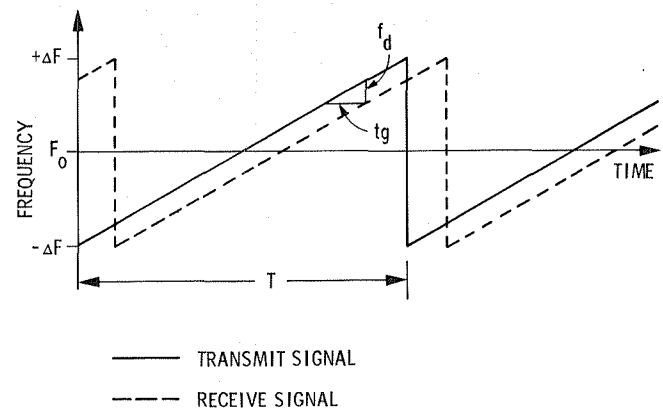


Fig. 6. FM/CW signal relationship

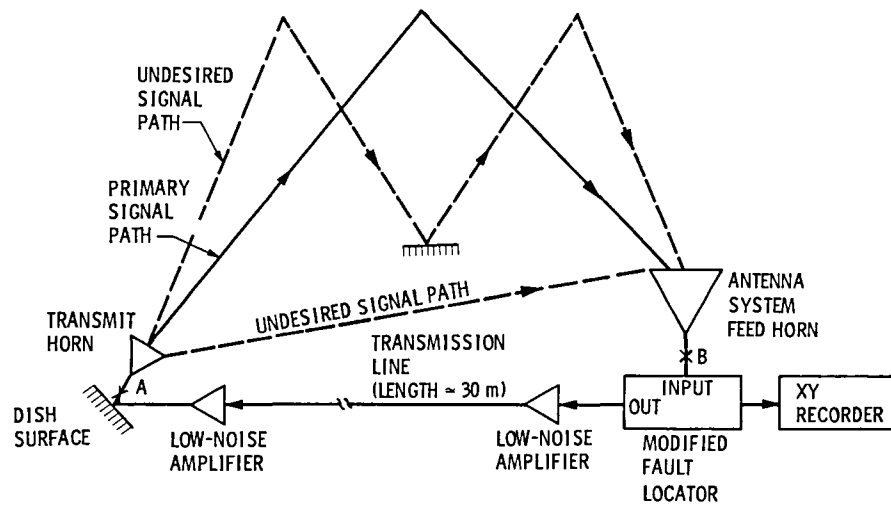


Fig 7. Block diagram of test setup for measurement of primary and multipath signal delays from Points A to B

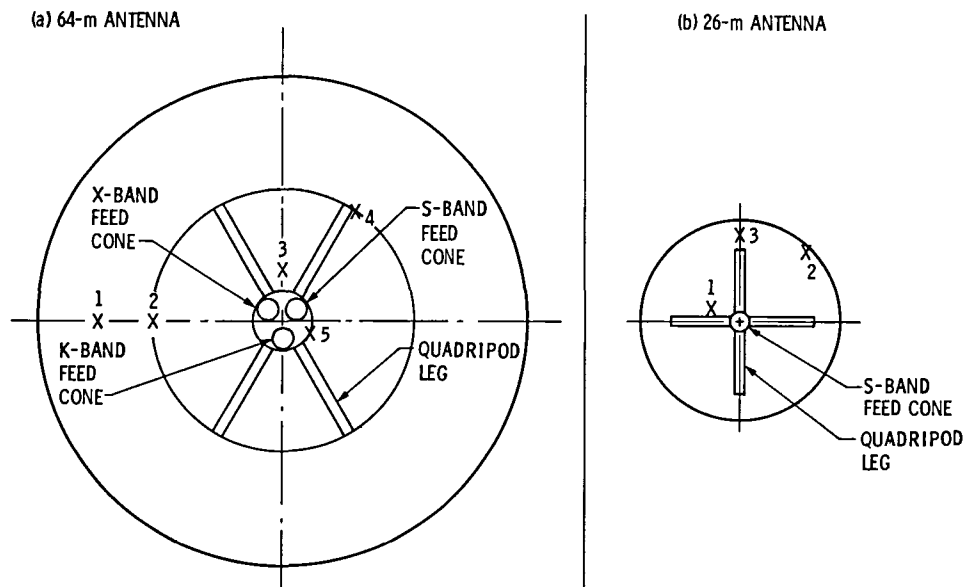


Fig. 8 Transmit horn locations for tests on (a) 64-m antenna and (b) 26-m antenna

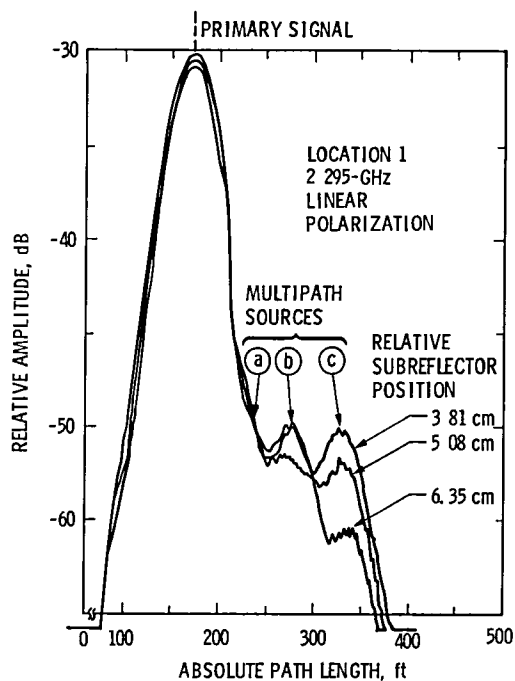


Fig 9 Sample recording of path lengths and relative signal magnitudes of primary and secondary signals for a 64-m antenna test setup The instrument records path length in units of feet and therefore it is not meaningful to convert the horizontal scale to metric units.

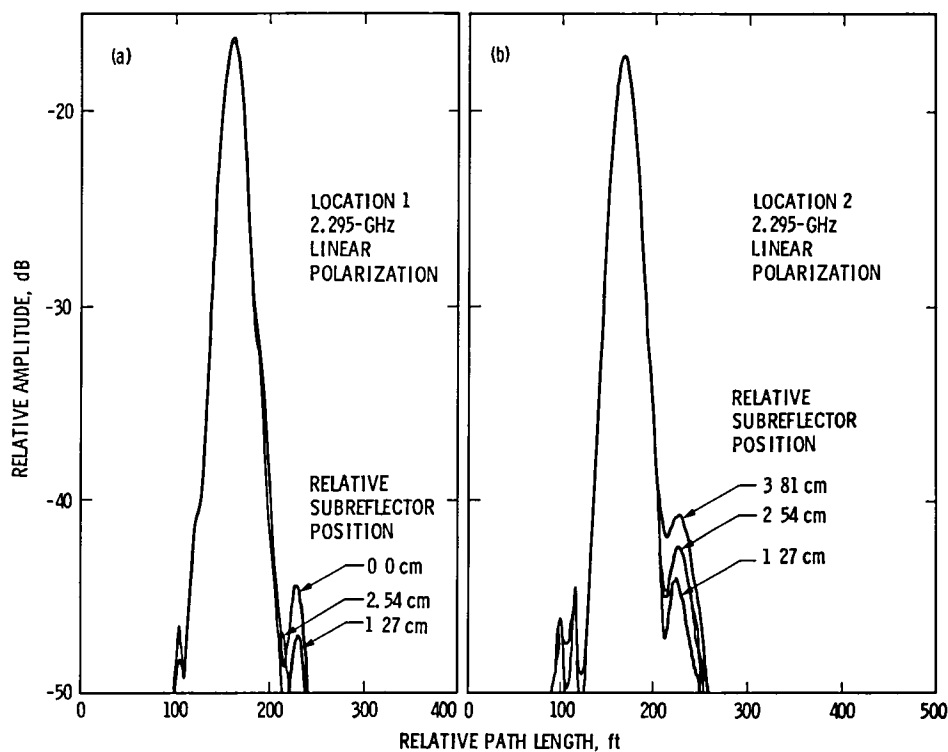


Fig. 10. Sample recordings of path lengths and signal magnitudes of primary and secondary signals for a 26-m antenna test setup with the transmit horn located (a) close to side of cone housing and (b) near the edge of paraboloid To convert horizontal scale to absolute path length for either plot, subtract the reference path length measurement of 95 ft

An M -ary Coherent Optical Receiver for the Free-Space Channel

V A Vilnrotter

Telecommunications Systems Section

The free-space channel is an ideal medium for communicating by means of spatially and temporally coherent optical fields. Here we derive the structure of a coherent optical receiver for M -ary signals assuming Poisson detection statistics. Receiver performance is evaluated for an M -ary signal set consisting of orthogonal Walsh-functions under the assumption of high-intensity symbol counts. The asymptotic performance bound is examined in the limit as both the dimension of the signal-set and the receiver bandwidth become arbitrarily large and it is shown that on the average a maximum of nearly three bits of information can be encoded onto each received photon using the above modulation scheme.

I. Introduction

In recent years, considerable effort has been devoted to the problem of deep-space communication by means of modulated optical fields (Refs. 1,2). Due to the inherently high gain of diffraction-limited optical antennas, optical communication systems offer the potential for high data-rate communication over interplanetary distances. However, since optical fields are susceptible to attenuation and distortion by the terrestrial atmosphere, current efforts have concentrated on communication systems operating over the free-space channel, which provides a distortionless medium for the propagation of optical fields. This type of free-space system could be implemented by placing the optical receiver outside of the terrestrial atmosphere (i.e., in Earth orbit) and relaying the received data

to ground by means of weather-independent RF data-links. In space, the relay receiver observes undistorted optical fields, and can therefore be designed to take advantage of the spatial and temporal coherence of the received field. In the following sections, we derive the structure and evaluate the performance of an optical receiver that coherently combines the received field with a locally generated optical field prior to photodetection, and attempts to retrieve the transmitted data by processing the detector's response to the combined field. We develop a mathematical model for the detection process, derive the structure of a maximum a posteriori (MAP) decoder and evaluate receiver performance for the case of orthogonal M -ary signals based on the well-known Walsh functions, a signaling scheme that allows for considerable simplifications in the subsequent analysis.

II. Coherent Optical Receiver Model

The coherent optical receiver considered here adds a strong, locally generated field coherently to the received optical field prior to photodetection, as shown in Fig 1. The received field can be modelled as a plane-wave

$$U_R(t) = U_R m(t) \exp [j (2\pi\nu t + \phi_R(t))] \quad (1)$$

where $2\pi\nu$ is the radian frequency of the received field, U_R is the (real) field amplitude, $m(t)$ is the modulation ($|m(t)| \leq 1$), and $\phi_R(t)$ is a random phase process associated with the received field. Similarly, we model the local field as

$$U_L(t) = U_L \exp [j (2\pi\nu t + \phi_L(t))] \quad (2)$$

where now U_L is an equivalent (real) local field amplitude referred to the receiver aperture, and $\phi_L(t)$ is again a random phase process. The detector responds to the instantaneous power of the sum field $s(t)$ by releasing free electrons from its active surface. If the detector bandwidth is sufficiently great, then for a given modulation function the detector output process $x(t)$ can be modelled as an inhomogeneous Poisson process, represented by a sequence of randomly occurring impulses. If the phase of the local field can be made to match that of the received field ($\phi_R(t) \cong \phi_L(t)$), then the modulation function $m(t)$ gives rise to an intensity function $\lambda(t)$, where

$$\begin{aligned} \lambda(t) &= \frac{\eta}{h\nu} A_R |s(t)|^2 \\ &\cong \frac{\eta}{h\nu} A_R (U_L^2 + 2 U_L U_R m(t)) \triangleq \lambda_L + \lambda_s(t) \end{aligned} \quad (3)$$

where we have assumed that $U_L \gg U_R$. (In Eq (3) η is the detector quantum efficiency, h is Planck's constant, ν is the frequency of the optical field and A_R is the area of the collecting aperture.) The count intensity therefore contains a constant term due to the local field and a modulation-dependent term due to the crossproduct of the local and received fields

Let the duration of each symbol be T seconds, and consider the time interval $[0, T)$. Assuming an infinite bandwidth detector, the output process can be represented as a sequence of randomly occurring impulses. With k such impulses occurring in $[0, T)$, a particular sample function of the output process can be represented as

$$x(t) = \begin{cases} 0 & , k = 0 \\ \sum_{n=1}^k \delta(t - t_n) & , k \geq 1 \end{cases} \quad (4)$$

Following Snyder (Ref 3) we define the "sample function density" for this inhomogeneous Poisson process as

$$p(x(t), 0 \leq t < T) = \begin{cases} \exp \left[- \int_0^T \lambda(t) dt \right] & , k = 0 \\ \left(\prod_{n=1}^k \lambda(t_n) \right) \exp \left[- \int_0^T \lambda(t) dt \right] & , k \geq 1 \end{cases} \quad (5a)$$

or equivalently

$$p(x(t), 0 \leq t < T) = \exp \left\{ - \int_0^T [\lambda(t) - x(t) \ln \lambda(t)] dt \right\} \quad (5b)$$

where Eq (5b) follows from the sifting property of delta functions. Note that for $k = 0$, $x(t) = 0$ in $[0, T)$, and therefore (5b) reduces to

$$\exp \left\{ - \int_0^T \lambda(t) dt \right\}$$

as in (5a).

III. Decoding of M -ary Coherent Signals

Consider the case where the transmitter generates one of M symbols in the time interval $[0, T)$. If the i th symbol is transmitted, the modulation function $m_i(t)$ generates the i th Poisson intensity $\lambda_i(t)$. Under the maximum a posteriori (MAP) decoding criterion the decoder selects that symbol whose probability, conditioned on the received sample function $x(t)$, is the greatest. Denoting the i th hypothesis by H_i , this requires computing $\text{Prob}(H_i | x(t), 0 \leq t < T)$ for all i , and selecting the hypothesis corresponding to the greatest posterior probability. Letting $P(H_i)$ denote the a priori probability of H_i , and $p(x(t), 0 \leq t < T | H_i)$ denote the conditional sample

function density, the decoder can equivalently compute

$$P(H_i) p(x(t), 0 \leq t < T | H_i) / p(x(t), 0 \leq t < T)$$

Keeping only hypothesis-dependent terms and taking the natural log, the decoder evaluates

$$\Lambda_i = Y_i + \int_0^T x(t) \ln \left(1 + \frac{\lambda_{si}(t)}{\lambda_L} \right) dt \quad (6)$$

where

$$Y_i = \ln P(H_i) - \int_0^T \lambda_{si}(t) dt$$

Based on Eq (6), the MAP decoder for M -ary signals can be implemented as shown in Fig. 2. Upon observing a particular sample function of the detector output process $x(t)$, the decoder multiplies this output by a set of hypothesis-dependent logarithmic weighting functions, integrates for T -seconds, adds the appropriate hypothesis-dependent bias terms and selects the hypothesis corresponding to the largest resulting value

IV. Receiver Performance

The receiver structure derived in the previous section applies to a general class of M -ary symbols, since no restrictions were imposed on the modulating functions (other than the normalization condition $|m(t)| \leq 1$). However, it is difficult to evaluate receiver performance for arbitrary modulation formats and a priori probabilities. The analysis may be greatly simplified for the case of equilikely orthogonal symbols generated by modulating functions that assume only a few discrete values. For purposes of analysis, a particularly simple signal set consists of the set of orthogonal Walsh functions $\{wal(i, t)\}$ which are defined and discussed in Appendix A. Each function assumes the values ± 1 in a given T -second interval (We exclude the zeroth Walsh function $wal(0, t)$ from the signal set because it is dissimilar from all other Walsh elements in that its average value is not zero, and would therefore require special treatment in the analysis). For equilikely Walsh symbols we let $P(H_i) = 1/M$ and $m_i(t) = wal(i, t)$, in which case the i th Poisson intensity function becomes $\lambda_i(t) = \lambda_L + \lambda_s wal(i, t)$, $i = 1, 2, \dots, M$. Now the decoder computes only

$$\Lambda_i = \int_0^T x(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L} wal(i, t) \right) dt \quad (7)$$

because the additive bias no longer depends on which hypothesis occurred

Since $x(t)$ is a sample function of a random process, its weighted integrals Λ_i are random variables. The second characteristic function of these random variables may take on either of two forms, depending on the relation between the transmitted symbol and the weighting applied to the detector output. If H_i is true, then let $\psi_{ii}(\omega)$ denote the second characteristic function of Λ_i , and let $\psi_{ik}(\omega)$ denote the second characteristic function of Λ_k , $k \neq i$. In order to determine $\psi_{ii}(\omega)$, we note from the properties of Walsh functions that $x(t)$ is a Poisson process with rate $(\lambda_L + \lambda_s)$ for a total of $T/2$ seconds, during which time a weighting factor of

$$\ln(1 + \lambda_s/\lambda_L)$$

is applied, while for the remaining $T/2$ seconds the rate parameter and weighting factor become $(\lambda_L - \lambda_s)$ and

$$\ln(1 - \lambda_s/\lambda_L)$$

respectively. Since intervals with different weighting factors are disjoint, integration over these regions yields independent weighted Poisson random variables. Hence, under H_i we can equivalently express Λ_i as

$$H_i, \Lambda_i = \int_0^{T/2} x^+(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) dt + \int_{T/2}^T x^-(t) \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) dt \quad (8)$$

where $x^+(t)$ is a Poisson process with intensity $(\lambda_L + \lambda_s)$, and $x^-(t)$ is a Poisson process with intensity $(\lambda_L - \lambda_s)$. The characteristic function for Λ_i under hypothesis i therefore becomes the product of the characteristic functions corresponding to each integral in (8), hence the second characteristic function can be expressed as

$$\begin{aligned} \psi_{ii}(\omega) &= \ln E \{ \exp(j\omega \Lambda_i) \} \\ &= \frac{T}{2} \lambda_L \left[\exp \left(j\omega \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) \right) + \exp \left(j\omega \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) \right) - 2 \right] \\ &\quad + \frac{T}{2} \lambda_s \left[\exp \left(j\omega \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) \right) - \exp \left(j\omega \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) \right) \right] \end{aligned} \quad (9)$$

A similar argument can be used to derive $\psi_{i\ell}(\omega)$. For this case the Poisson intensity is still defined by the i th Walsh function, but now the weighting depends on the ℓ th Walsh function, $\ell \neq i$. Now Λ_ℓ is obtained by integrating $x(t)$ over four disjoint intervals of equal duration during which all four possible combinations of count intensities and weights are assumed

$$\begin{aligned} H_i \Lambda_\ell = & \int_0^{T/4} x^+(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) dt \\ & + \int_{T/4}^{T/2} x^+(t) \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) dt \\ & + \int_{T/2}^{3T/4} x^-(t) \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) dt \\ & + \int_{3T/4}^T x^-(t) \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) dt \end{aligned} \quad (10)$$

In this case $\psi_{i\ell}$ reduces to

$$\begin{aligned} \psi_{i\ell}(\omega) = & \frac{T}{2} \lambda_L \left[\exp \left(j\omega \ln \left(1 + \frac{\lambda_s}{\lambda_L} \right) \right) \right. \\ & \left. + \exp \left(j\omega \ln \left(1 - \frac{\lambda_s}{\lambda_L} \right) \right) - 2 \right] \end{aligned} \quad (11)$$

In typical applications, the total power of the local field is great enough to generate a high intensity process at the detector output. Expanding the exponentials in (9) and (11) and letting the intensity due to the local field become arbitrarily large ($\lambda_L \rightarrow \infty$) we obtain the limiting forms

$$\psi_{11}(\omega) \xrightarrow{\lambda_L \rightarrow \infty} j\omega(2K_s) - \frac{\omega^2}{2}(4K_s) \quad (12a)$$

$$\psi_{i1}(\omega) \xrightarrow{\lambda_L \rightarrow \infty} j\omega(-2K_s) - \frac{\omega^2}{2}(4K_s) \quad (12b)$$

where we define K_s as the average symbol count that would be generated by the received field if it were direct-detected by the same receiver.

$$K_s = \frac{\eta}{h\nu} A_R T U_R^2$$

Equation (12a) is recognized as the second characteristic function of a Gaussian random variable with mean value $(4K_s)$ and variance $(4K_s)$, while (12b) corresponds to a Gaussian random variable with zero mean and variance $(4K_s)$. It follows that as the average intensity λ_L becomes suitably great, the random variables Λ_i may be approximated by Gaussian random variables with mean and variance as defined above.

The probability of correct decoding can be found by computing the probability that given H_i , Λ_i exceeds all other Λ_ℓ , $\ell \neq i$. Making use of the Gaussian approximation, the probability of correct decoding, $P(C)$, can be expressed as

$$\begin{aligned} P(C) = & \int_{-\infty}^{\infty} dx \frac{e^{-(x - 4K_s)^2 / 2(4K_s)}}{\sqrt{2\pi(4K_s)}} \times \\ & \left[\int_{-\infty}^x dy \frac{e^{-y^2 / 2(4K_s)}}{\sqrt{2\pi(4K_s)}} \right]^{(M-1)} \end{aligned} \quad (13)$$

while the symbol error probability becomes $P(E) = 1 - P(C)$. Equation (13) is well known in the context of M -ary orthogonal signal detection in additive Gaussian noise, a similarity that stems from the Gaussian approximation made in Eq. (12). Assuming a photodetector quantum efficiency of one ($\eta = 1$), it is convenient to define the "photon information rate" as $\rho \triangleq (\log_2 M)/K_s$, which can be interpreted as the average information in bits encoded onto each detected photon (Ref. 4). Since each symbol gives rise to $(M/2)/(M-1)$ bit errors, the average bit error probability $P_B(E)$ can be expressed as

$$\begin{aligned} P_B(E) = & \left(\frac{M/2}{M-1} \right) \left\{ 1 - \int_{-\infty}^{\infty} dz \frac{e^{-z^2}}{\sqrt{\pi}} \times \right. \\ & \left. \left[1 - \frac{1}{2} \operatorname{Erfc} \left(z + \sqrt{\frac{2 \log_2 M}{\rho}} \right) \right]^{(M-1)} \right\} \end{aligned} \quad (14)$$

where we recognized the relation between the bracketed term in (13) and the complementary error function. Equation (14) has been tabulated extensively in the literature (Refs. 5, 6). Figure 3 shows graphs of the average bit-error probability $P_B(E)$ as a function of ρ for increasing M , obtained from the tabulated values in (Ref. 6). Note that photon information

errata
 3-9-82

rates of roughly 1 bit per photon can be achieved in the range of bit error probabilities of interest ($P_B(E) \lesssim 10^{-3}$) with signal sets of moderate dimension ($2^5 < M < 2^{10}$). The limiting behavior of $P_B(E)$ as the number of transmitted symbols becomes arbitrarily large can be obtained from the corresponding results derived for additive Gaussian noise by a simple transformation of variables (Refs 5, 6)

$$\lim_{M \rightarrow \infty} P_B(E) = \begin{cases} \frac{1}{2}, & \frac{2}{\rho} < \ln 2 \\ 0, & \frac{2}{\rho} > \ln 2 \end{cases} \quad (15)$$

It follows that arbitrarily low bit error probabilities can be achieved as M approaches infinity, as long as the inequality $\rho < 2/\ln 2 \cong 2.89$ bits/photon is satisfied. This implies that when operating with wideband, high quantum-efficiency photodetectors, the M -ary coherent optical system described above can theoretically transfer several bits of information per received photon on the average (depending on the required bit error probability), but the limiting value of 2.89 bits/photon cannot be exceeded, regardless of system complexity.

V. Conclusions

A coherent optical receiver model for M -ary orthogonal signals has been examined. The structure of the MAP decoder designed for arbitrary signals was developed, and receiver performance evaluated for the special case of M -ary orthogonal signals derived from Walsh functions. This specialization enabled the development of a mathematically rigorous solution for receiver performance in the limiting case of high intensity detection, where a Gaussian approximation could be invoked. It should be emphasized, however, that these results also apply to a more general class of orthogonal signals, as long as the Gaussian approximation for the test random variables $\{\Lambda_i\}$ can be justified. Receiver performance was evaluated in terms of the "photon information rate" ρ , which is a measure of the average information that can be encoded onto each detected photon. The equivalent bit error probability depends both on ρ and on the signal-set dimension M . It was found that reliable communication could not be achieved at photon information rates exceeding 2.89 bits/photon. However, reliable communication at photon information rates exceeding one bit per photon appears feasible with M -ary coherent optical systems employing high quantum-efficiency, wideband photodetectors, if negligibly small optical phase error can be maintained. The effects of phase error and of other external disturbances (such as background radiation) on the performance of M -ary coherent optical receivers remains to be examined in future studies.

References

1. Vlnrotter, V. A., and Gagliardi, R. M., "Optical Communications Systems for Deep-Space Applications," Publication 80-7, Jet Propulsion Laboratory, Pasadena, Calif., Mar. 15, 1980.
2. Gagliardi, R. M., Vlnrotter, V. A., and Dolinar, S. J., "Optical Deep Space Communication via Relay Satellite," Publication 81-40, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.
3. Snyder, D. L., *Random Point Processes*, J. Wiley, New York, N.Y., 1975.
4. McEliece, R. J., and Welch, L. R., "Coding for Optical Channels with Photon Counting," in *The DSN Progress Report 42-52*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1979.
5. Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill, New York, N.Y., 1966.
6. Lindsey, W. C., and Simon, M. K., *Telecommunication Systems Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1973.
7. Harmuth, F. H., *Transmission of Information by Orthogonal Functions*, Springer-Verlag, Berlin, Germany, 1969.

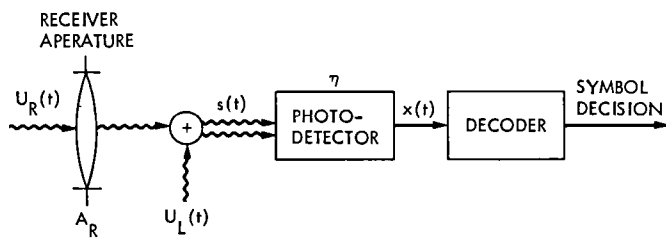


Fig 1 Coherent optical receiver block diagram

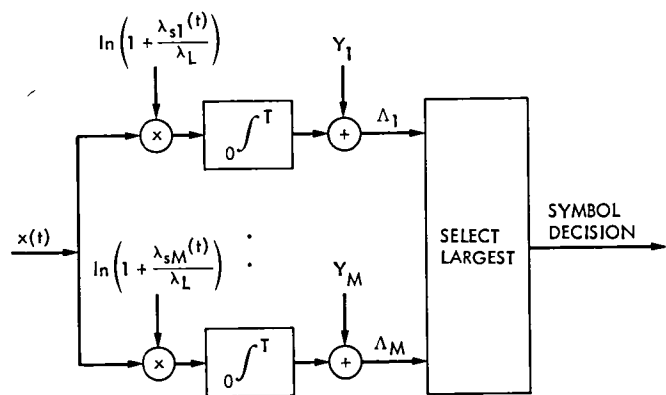


Fig. 2. MAP receiver structure

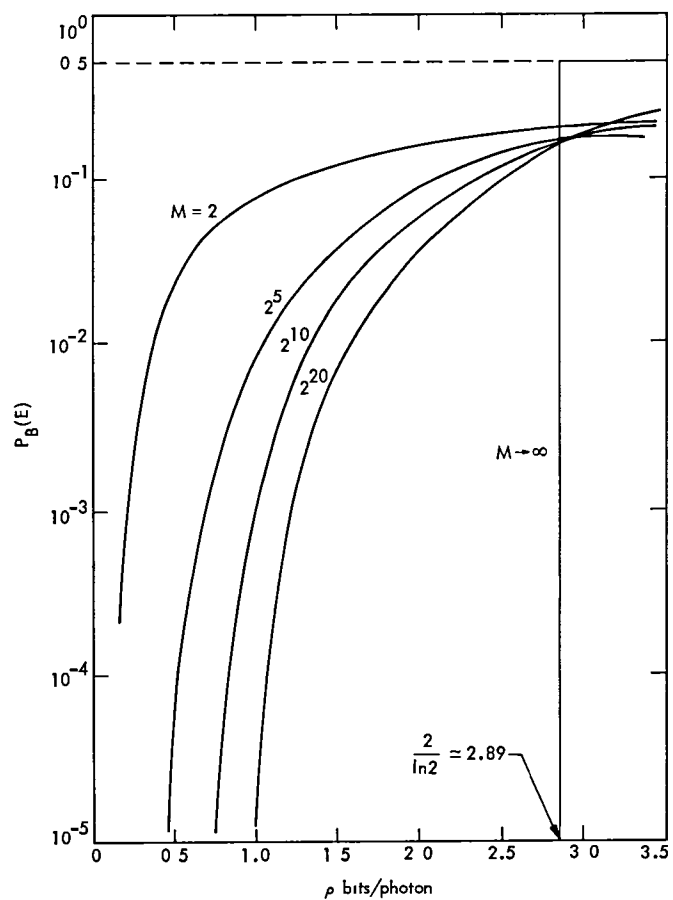


Fig. 3. M -ary coherent optical receiver performance: $P_B(E)$ vs ρ for various M

Appendix A

Properties of Walsh Functions

An extensive discussion of the properties of Walsh functions can be found in Ref 7. For our interest here, it is convenient to define the Walsh functions $wal(i, t)$ over the interval $[0, T)$, indexed by i . The Walsh functions can be generated by the equation

$$wal(2m + p, t) = (-1)^{\lfloor m/2 \rfloor + p} \left\{ wal(m, 2t) + (-1)^{m+p} wal\left(m, 2\left(t - \frac{T}{2}\right)\right) \right\} \quad (A-1)$$

$p = 0$ or $1, m = 0, 1, 2,$
where

$$wal(0, t) = \begin{cases} 1, & 0 \leq t < T \\ 0, & \text{elsewhere} \end{cases} \quad (A-2)$$

In Eq (A-1) $\lfloor m/2 \rfloor$ means "greatest integer less than or equal to $m/2$ ". As an example, let us generate $wal(1, t)$. This requires setting $m = 0, p = 1$. Hence

$$wal(1, t) = (-1) \left\{ wal(0, 2t) - wal\left(0, 2\left(t - \frac{T}{2}\right)\right) \right\} \quad (A-3)$$

All subsequent elements are obtained from previous elements by applying scaling, shifting and sign reversal operations

The product of two Walsh functions is another Walsh function.

$$wal(i, t) wal(\ell, t) = wal(i \oplus \ell, t) \quad (A-4)$$

where \oplus stands for modulo 2 addition of the indices expressed in binary form. Note that the product of a Walsh function with itself always yields $wal(0, t)$.

Walsh functions are orthogonal and therefore obey the relation

$$\frac{1}{T} \int_0^T wal(i, t) wal(\ell, t) dt = \delta_{i\ell} \quad (A-5)$$

where $\delta_{i\ell}$ is the Kroenecker delta. Since each Walsh function $wal(i, t), i \geq 1$ is orthogonal to $wal(0, t)$, it follows that

$$\int_0^T wal(i, t) dt = 0, i \geq 1 \quad (A-6)$$

Therefore each of these functions assume the values $+1$ and -1 for exactly half the total duration, or $T/2$ seconds

Consider any two Walsh functions $wal(i, t)$ and $wal(\ell, t)$ satisfying $i, \ell \geq 1, i < \ell$. We maintain that a point-by-point comparison yields all four possible combinations of $+1$ and -1 for exactly one quarter of the total duration, or $T/4$ seconds. This assertion can be proven as follows: the product of any two Walsh functions is another Walsh function, which must assume the values $+1$ and -1 for equal lengths of time in order to satisfy (A-6). The two Walsh components of the product must therefore agree in sign for $T/2$ seconds (to generate $+1$ in the product) and disagree in sign for $T/2$ seconds (to generate -1 in the product). Let d_{++} denote the total duration over which both component Walsh functions agree in sign and are positive, d_{--} denote the duration over which they agree in sign and are negative, and likewise d_{+-} and d_{-+} denote durations over which the component functions disagree in sign. (Let the first sign in the subscript refer to the Walsh function with the lower index, i .) Clearly, $d_{++} + d_{--} = d_{+-} + d_{-+} = T/2$ in order for the product to satisfy (A-6). It remains to be shown that

$$d_{++} = d_{--} = d_{+-} = d_{-+} = T/4 \quad (A-7)$$

Let $d_{++} = d_{--} + \Delta$ for some Δ . Then it must be true that $d_{-+} = d_{+-} + \Delta$ in order for $wal(i, t)$ to satisfy (A-6), but we must also have $d_{+-} = d_{-+} + \Delta$ since $wal(\ell, t)$ must also satisfy (A-6). Hence $\Delta = 0$, and $d_{++} = d_{--}$. A parallel argument shows that $d_{-+} = d_{+-}$, and Eq (A-7) follows

Thermal Measurement Technique of Rib Elements on DSN Antenna Structure

S Glazer

Applied Mechanics Technology Section

G Gale

DSN Engineering Section

Thermal gradients on large DSN antennas may cause small structural deflections which affect the RF gain. These deflections may become very important at the high frequencies encountered in the Ka band. This report discusses an experimental test method employing an imaging infrared radiometer, which promises to yield accurate thermal data for these large antennas in a timely and convenient manner. The ultimate goal is to use the data in structural analysis codes to determine the effects on dish deflection and RF performance. Basic radiometric theory is presented, along with a description of the test apparatus and current experimental test procedures. A sample set of antenna rib temperatures using the developed methodology is also presented.

I. Introduction

Changes in length of rib elements on DSN antennas may cause the surface shapes to change enough to affect the RF gain. This is especially true if the antenna is to be used in the Ka band. A change in surface shape by as little as 1/50 of the wavelength (0.2 mm) will affect the gain. Complex shadow and illumination patterns resulting whenever the antenna does not track the sun may cause local rib temperature variations, which may in turn induce member strains. These complex thermal patterns are further affected by the transient tracking path, daily and seasonal weather conditions, and antenna geometry. This report describes an infrared radiometric technique for measuring antenna rib structure temperatures with-

out attaching thermometers or thermocouples to the ribs. The data will subsequently be used as input to structural analysis codes to determine antenna surface deflections.

Several years ago an attempt was made to measure thermal gradients within the rib structure by attaching many thermocouples at the joint points and midway along the rib members. Data obtained from these measurements were inconsistent and data reduction was abandoned.

The present method employs a real-time infrared radiometer operated in the near infrared region to view the thermal emissions from the ribs. The IR signal is recorded on a standard video tape recorder, allowing a large quantity of data

to be obtained in a short period of time, and subsequent off-line data reduction. The advantages of this method are that thermal changes during the data taking period are held to a minimum, hardware which could interfere with normal use need not be attached to the antenna, and data analysis may be accomplished at a later, more convenient time.

II. Description of Infrared Radiometry

This section presents some of the background required for an understanding of the basic infrared radiometric method. This background includes a brief discussion of the fundamentals of radiation heat transfer, general principles of radiometer operation, and some of the factors which affect instrument accuracy.

A. Fundamentals of Radiation Heat Transfer

All objects whose temperatures are above absolute zero emit energy by the mechanism of radiation. This energy may be considered as being emitted in discrete packets sometimes called photons. Each photon also has an associated wavelength. A "black body" is defined as an object which emits the maximum possible theoretical energy over all wavelengths. The spectral radiant emittance is defined as the theoretical maximum energy flux which may be emitted at any wavelength λ by a black body at temperature T . The expression for the spectral radiant emittance is

$$\text{Spectral Radiant Emittance } e = \frac{3.75 \times 10^{-12}}{\lambda^5 (e^{1438/\lambda T} - 1)} \frac{W}{\text{cm}^2 \cdot \mu\text{m}} \quad (1)$$

where

λ = wavelength in cm

T = temperature in K

Figure 1 graphically displays the spectral radiant emittance as a function of wavelength for black bodies of various temperatures. For many practical purposes, the sun may be considered a black body radiating at approximately 6000 K. As may be seen, the wavelength of maximum energy emission is approximately $0.5 \mu\text{m}$, conveniently well within the narrow wavelength band to which our eyes are sensitive (the visible portion of the electromagnetic spectrum ranges from about $0.4 \mu\text{m}$, to $0.7 \mu\text{m}$). Cooler black objects emit less energy in the visible portion of the spectrum. In fact, as black body temperature decreases, the wavelength of peak energy emission increases. As seen in Fig 1, black objects at room temperature emit nearly all their energy between about 4 and $30 \mu\text{m}$, peaking at approximately $10 \mu\text{m}$. Our eyes are therefore not sensitive to the energy emitted by room temperature objects,

rather, they may be seen only by virtue of shorter wavelength energy (as for example, from the sun or artificial light sources) reflected from the surface. Infrared radiometers are generally sensitive to energy of wavelengths longer than $0.7 \mu\text{m}$ (infrared, or beyond the red, wavelengths).

As will be discussed shortly, the infrared radiometer used in the present study belongs to an important class of instruments which are sensitive not to the total energy emitted, but rather to the total number of photons striking the detector element. If we redefine the spectral radiant emittance as the maximum theoretical photon flux at any value of wavelength, then the following expression may be developed

$$\text{Spectral radiant emittance } p = \frac{1.89 \times 10^7}{\lambda^4 (e^{1438/\lambda T} - 1)} \frac{\text{photons}}{\text{sec} \cdot \text{cm}^2 \cdot \mu\text{m}} \quad (2)$$

where

λ = wavelength in cm

T = temperature in K

Figure 2 is a plot of this parameter as a function of wavelength for several black bodies of various temperatures.

In practice, black bodies do not exist. If emissivity is defined as the ratio of actual energy flux (or photon flux) emission per unit wavelength to the maximum theoretical (black body) rate, the emissivity of various surfaces would be found to depend on the wavelength, surface temperature, and direction of emission. For many real surfaces, the diffuse grey body approximation is made. A grey body is assumed to emit a constant fraction of the maximum theoretical energy at each wavelength. Further, the energy emitted by a diffuse grey body per unit projected surface area is assumed independent of direction. Emissivity defined in this way is often called the total hemispherical emissivity. Frequently, thermal analysis requires the dual grey body approximation, where two values of hemispherical (average) emissivity are used, one value for shorter wavelengths where the solar energy and photon emission peaks ($< \sim 4 \mu\text{m}$) and one value for longer wavelengths, where nearly all the energy and photon emission from room temperature objects occur ($> \sim 4 \mu\text{m}$). The former quantity is usually referred to as the solar absorptivity (α_s), and the latter simply as the emissivity (ϵ).

Energy and photons incident upon a surface may be absorbed, reflected, or transmitted through the surface. In general, each of these quantities is a function of wavelength, surface temperature, and direction. It may be shown that in the steady state, the directional, spectral emissivity of a

surface at temperature T equals the directional, spectral absorptivity of that surface. This is known as Kirchoff's Law. If the following definitions are made

α = absorptivity = total fraction of incident energy absorbed

ρ = reflectivity = total fraction of incident energy reflected

τ = transmissivity = total fraction of incident energy transmitted through,

then it may be seen that

$$\alpha + \rho + \tau = 1 \quad (3)$$

If the surface is opaque to all radiation, then

$$\alpha + \rho = 1 \quad (4)$$

If Eq (1), the spectral radiant emittance for a black body, is integrated over all wavelengths, then the total energy flux (called the black body emissive power, e_{bb}) is

$$e_{bb} = \sigma T^4 \quad (5)$$

where σ is the Stefan-Boltzmann constant. For a real room temperature body assumed to be diffuse having a total hemispherical emissivity ϵ , the emissive power is

$$e = \epsilon \sigma T^4 \quad (6)$$

Thus the total energy leaving a surface per unit time and surface area is defined as the radiosity R , as

$$R = \epsilon \sigma T^4 + \frac{(\text{incident energy})}{\text{unit-time-area}} \times \rho \quad (7)$$

Kirchoff's law in conjunction with Eq (4) for an opaque surface indicates that Eq (7) may be rewritten as

$$R = \epsilon \sigma T^4 + \left(\frac{\text{incident energy}}{\text{unit time-area}} \right) \times (1 - \epsilon) \quad (8)$$

A relation similar to Eq. (8) may be derived for the photon flux radiosity. Equation (8) states that the energy (and number of photons) leaving a surface is composed of two parts: that emitted and that reflected. The closer the surface emissivity to 1.0, the smaller the relative importance of the reflected component. Surfaces of many organic materials,

including paints, have emissivities in the 0.8-0.95 range. A more complete discussion of radiation heat transfer may be found, for example, in Ref. 1.

B. Principles of Radiometer Operation

Radiometers may be divided into two general classes: those which sense the energy emitted (plus reflected), called thermal detectors, and those which detect the total number of photons emitted (and reflected), or photon detectors. Thermal detectors may employ the bolometric effect, as do thermistors and platinum RTD's, the thermovoltaic effect, (such as thermocouples and thermopiles), thermopneumatic or pyroelectric effects. Photon detectors may use the photovoltaic effect, as do solar cells, the photoconductive effect, or photoemissive effect.

The infrared radiometer used in the present analysis is a photon-sensitive device which utilizes the photoconductive effect. The electrical conductivity of this type of detector is a strong function of the photon flux. Since the detector does not respond to a change in its temperature, its speed of response is very fast, on the order of μsec . This type of detector is thus well suited to the rapid scene scanning necessitated by video compatible scan rates.

Single point sensing (nonscanning) and single horizontal line scanning infrared radiometers exist and are useful in certain applications. The most general and useful system is of the two-dimensional imaging type. Typically, photons from the scene are focused onto the small fixed detector element by means of rotating mirrors. The solid angle subtended by the detector is called the Instantaneous Field of View (IFOV), and is the smallest physical unit into which the scene may be resolved. It is also referred to as a resolution element. Photons from the overall scene being viewed, or Field of View (FOV), are focused sequentially onto the detector, resolution element by resolution element, until the entire scene has been scanned. After reference to some specified level, pointwise scene radiosities are displayed in a suitable manner, usually on a video monitor. In this format of data presentation, radiosity levels are usually indicated by local display brightness, with higher levels often displayed as brighter regions. Other modes of data presentation available with the specific instrument being used are discussed in Section III.

C. Factors Affecting Accuracy

Modern infrared radiometers infer object temperatures from radiosity levels observed. As previously discussed, radiosity, which is the total energy (or photon) flux leaving an object's surface, is composed of emitted plus reflected components. Only the emitted energy (or photon flux) component is meaningful in determining object temperature. It is

therefore necessary to separate the total radiosity into its constituents. In addition, infrared radiometer operation in an environment with absorbing and scattering media between the target and the instrument can further attenuate the radiosities measured. Finally, a large number of instrument-dependent properties, such as optical distortions, nonlinearity of detector signal amplification, electronic and radiometric noise, and detector and electronics transient response affect instrument accuracy. A brief discussion of the importance of some of these factors is included here. More complete treatments of this topic may be found in, for example, Refs. 2 and 3.

As previously discussed, all nonblack objects reflect some fraction of the energy incident upon them by other objects. The amount of energy incident upon the target (object whose temperature is to be measured) depends on (1) the temperature of the surrounding objects, (2) the view that the target surface has to each of the surrounding objects (or view factors), (3) diffusivity and specularity of all surfaces, (4) the emissivities of all the surrounding surfaces, and (5) the nature and quantity of all external energy (or photon) flux incident on all surfaces, i.e., the incident solar flux. In theory, all of the above factors must be known before target temperatures may be established with accuracy. Fortunately, physical test conditions of the present application, i.e., radiometric determination of thermal gradients on DSN antennas, permit use of many simplifying assumptions which eliminate the need for much of this detailed information. Following is a brief discussion of the simplifying assumptions being considered at this time in the final data reduction process.

The rib network comprising the antenna structure is highly complex and irregular. Any point on any rib member may have a physical view to several other rib elements, to the back of the dish surface, to the sky, and to the ground. All of these affect the incident energy and photon flux onto the target rib element. For the purpose of this analysis, the target rib will be assumed to view only the sky and the ground, not other portions of the antenna. This step is taken for several reasons. First, Eq. (8) indicates that the importance of energy or photons incident from other objects decreases as the target emissivity in the long infrared region ($> \sim 4 \mu\text{m}$) approaches 1.0. The two types of white paint used on DSN antennas, Triangle #6 and Americoat #99, have emissivities in the 0.85-0.90 range, hence energy (or photons) from all other sources of approximately the same temperature contribute only 10-15% to the overall radiosity. Second, it is assumed that the view factor from the outward facing target rib elements to all other surrounding ribs is small in comparison with the total hemispherical view. It is therefore reasonable to assume that each target rib section views only the sky and the ground, each approximately equally. For simplicity, the ground and the sky may each be assumed "black," ($\epsilon \approx 1.0$)

and the ground temperature set equal to the ambient air temperature. Correlations exist (Ref. 4) for clear sky effective temperature. The black body approximations for the sky and ground imply totally diffuse surfaces. Similarly, paints of the type used on the antennas typically are highly diffuse, and exhibit very small specular components. This assumption will be tested in the laboratory for the particular paints used. The use of these simplifications permits rib temperature determination without the necessity of complex thermal modeling.

The importance of reflected solar energy in the total radiosity measured will be ignored. Outside the earth's atmosphere, at a distance of 1 astronomical unit, the actual measured radiant emittance of solar radiation matches closely that of a black body at approximately 6000 K. However, the total energy flux through air mass 1 (the solar intensity measured with the sun directly overhead) is approximately 1000 W/m^2 . Beyond $7 \mu\text{m}$, only about 0.2% of this energy, or 2.0 W/m^2 is detected. By contrast, a black object at 30°C emits approximately 157 W/m^2 between 7 and $12 \mu\text{m}$. The diffusely reflected solar component from a target with an emissivity of 0.9 is thus less than 1.0% of the emitted energy in this wavelength band. For this reason, a spectral filter which prohibits energy and photons below $7 \mu\text{m}$ from passing will be used in conjunction with the infrared radiometer in future tests.

Transmission of energy through the earth's atmosphere is another factor affecting accuracy of radiometric measurements. In general, the mechanisms of absorption and scattering attenuate energy within the earth's atmosphere. The phenomenon of scattering causes collimated energy to be scattered diffusely. Effects of molecular scattering are negligible in the atmosphere for energy above about $2 \mu\text{m}$, whereas scattering from clouds, aerosols, and hazes is not and must be considered when they intervene between the radiometer and the target.

Absorption is highly significant in attenuating energy or photon fluxes in the earth's atmosphere. In the phenomenon of molecular and gaseous absorption, energy is absorbed and reemitted at the molecule or gas temperature. In the lower atmosphere, water vapor, carbon dioxide, ozone, and methane are the primary constituents causing significant energy attenuation. The extent of spectral attenuation is highly dependent on such factors as the local concentrations of these components, humidity, pressure, target to radiometer distance, and altitude. While spectral transmission through the earth's atmosphere is highly nonuniform and variable, it is generally restricted to the 3.5-5.0 μm and the 8.0-14 μm bands within the 2-20 μm thermal energy wavelength region. Typical atmospheric spectral transmission curves as a function of many of the parameters mentioned are presented in Ref. 2.

Additional factors affecting the accuracy of radiometric temperature determination involve instrument specific operating characteristics. These factors include the specific detectivity of the detector, which is basically the detector output signal-to-noise ratio at 1 watt of monochromatic input power, the responsivity, which is the detector gain per watt of input energy, amplifier linearity, electronic noise, optical aberrations, internal reference level stability with time, and the relationship between target size, scene scan rate, and overall system response. Thorough explanation of these and other factors is beyond the scope of the present report, and so will not be treated here. References 2 and 3, for example, provide adequate coverage of these topics. Only the first and last of the above factors will be very briefly discussed here, due to their particular importance in the present application.

The specific detectivity D^* depends principally on the spectral sensitivity of the detector material, and is thus basically a material property. The specific detectivity of a system may be enhanced by detector operation at cryogenic temperatures. Many modern radiometers thus employ liquid nitrogen cooled detectors. The specific detectivity is an important consideration when choosing the optimal detector element for a particular application. Generally, the highest values of D^* are desirable at the wavelengths where most of the target energy (or photons) is emitted. Figures 1 and 2 indicate that objects around room temperature (such as in the present application) emit the most energy at the longer wavelengths, so that the detector D^* should be optimized in this region for the present application.

Target size is also a very important consideration in infrared radiometry. As discussed, the instantaneous field of view (IFOV), or resolution element, is the smallest size target which may be resolved. However, due to overall instrument response time, a step change in target temperature takes a finite amount of time to be accurately indicated on the display. Since the IFOV is constantly moving in a scanning system, this translates into a finite distance from the start of a uniform temperature region. At video scan rates, a general rule of thumb is that uniform temperature targets should occupy at least 4-5 IFOV's for accurate temperature indication.

D. Minimum Resolvable Temperature Difference—An Accuracy Criterion

The minimum resolvable temperature difference (MRTD) is a commonly used accuracy parameter for infrared radiometers. This criterion uses four bar targets of fixed aspect ratio, but variable widths. In the MRTD test, the black target foreground and background are maintained at fixed and well-known temperature differences. The MRTD at any given target size (expressed in terms of bar cycles/milliradian of scene) is the minimum ΔT between target foreground and background at

which observers can distinguish the pattern through the electronic and radiometric noise. The MRTD is thus a curve which asymptotically approaches a single ΔT value. It is a useful tool in comparison of radiometric systems.

III. Description of the Infrared Camera Used

The infrared camera used is an Inframetrics Model 525 two-dimensional scanning radiometer. It uses a liquid-nitrogen-cooled HgCdTe detector element which is designed to be sensitive primarily to radiation in the 8-12 μm wavelength band. The system can measure temperatures over the range of -20 to $+1300^\circ\text{C}$, with a minimum resolvable temperature difference (MRTD) of 0.1°C for a black 30°C target which is at least 5 IFOV's in size. Thermal images are presented in real-time on a television monitor, with regions of higher radiosity nominally appearing brighter. System output is compatible with RS-330 video, enabling real-time recording of thermal imagery on a conventional video tape recorder. Subsequent quantification of the thermal data may be performed off line by replaying the video tape through a colorizing system which artificially assigns false colors to a maximum of six radiosity intensity bands. The same colorizing system may alternately be used in a black and white mode to artificially highlight all regions of the scene displaying the same selectable value of radiosity. It is this mode which permits most convenient scene temperature comparisons with a known reference. Further discussions of data analysis procedures are contained in Section IV.

The basic infrared radiometer without magnification optics has an instantaneous field of view (IFOV) of 2.0 mr, and an overall field of view (FOV) of $14 \times 18^\circ$. Guidelines discussed in Section II indicate that the smallest target for which temperatures may be accurately determined occupy 5 IFOV's. Without optical magnification, this corresponds to a minimum target of 18.0 inches at a nominal radiometer-to-antenna distance of 150 feet. A 10X magnifying telescope attachment is therefore used in the present application. This accessory narrows the overall FOV to $1.4 \times 1.8^\circ$, while reducing the IFOV to 0.2 mr. This causes a subsequent reduction in minimum target size to 1.8 inches at 150 feet, enabling accurate thermal resolution of the 2- to 3-inch-wide structural members. The 10X telescope has an 8-inch aperture, and uses both refractive and reflective optics. Descriptions of additional related equipment used in the thermal tests are provided in Section IV.

IV. Thermal Test Methodology

A. Initial Test Method

Throughout 1980 and 1981, a low level effort was conducted at the Goldstone DSCC to determine the feasibility

of infrared radiometry to yield meaningful antenna temperatures for subsequent use in structural analysis codes. Ground-based observations were made initially with the standard optics, and later, with a 3-power magnifier lens. These attempts proved fruitless, because the sizes of typical structural members on the Venus antenna (DSS 13) are too small for accurate thermal resolution from a distance of 150 feet, necessary for adequate ground-based spatial coverage. As discussed in Section II, a minimum of 5 IFOV's, or resolution elements, are required for accurate instrument response to a uniform temperature target. Further, the complex rib structure on the antenna made positive member identification with only the infrared image very difficult. Figure 3 is a wide angle picture, taken with a standard video camera, of the 32-m DSS 13. It illustrates the complexity of the rib structure. The subreflector and quadrupod legs may be seen through the perforated dish surface. These problems are magnified with the 64-m Mars antenna (DSS 14).

The next approach involved infrared inspection of structural members while based in a movable aerial work platform (a large cherry picker). Using this technique, the infrared instrument, operator, and platform operator conducted all operations from the cherry picker bucket, which was constantly maneuvered to within approximately 6 feet of the antenna structural members. This method yielded positive element identification and reasonable size views of much of the antenna. Figure 4 is an infrared image of a structural member taken from the cherry picker. This figure illustrates the iso-radiosity enhancement mode discussed in Section III.

The aerial survey technique included several serious drawbacks, however. Great care was required in platform maneuvering around the complex antenna structure, particularly during periods of antenna tracking. The requisite care in platform placement limited the speed of the overall data collection procedure to approximately one relatively complete thermal map every 3-4 hours. Contributing to the length was the requirement of moving the truck-based cherry picker to several positions around the antenna to obtain adequate coverage. Possible errors were introduced into the data gathered by this method by several factors. First, ambient conditions (i.e., air temperature, wind velocity, direction) could change appreciably over the several-hour test duration. Thus, initial data points could be taken under different conditions than those recorded near the test conclusion. Secondly, radiometer instrument internal reference level temperature may not have remained constant throughout the test. Typically, an external base level was established approximately once per hour by viewing a controlled black body reference source placed on the antenna. This may be inadequate and of questionable validity given the level of accuracy required.

B. Present Data Collection Method

The current test procedure again utilizes the infrared radiometer in a ground-based mode, but with a 10X magnifier lens. An adequate view of most of the antenna structure may be obtained from one, and at most two, fixed positions approximately 150 feet behind the antenna. The 10X telescope attachment, which is an off-the-shelf radiometer accessory, reduces the resolution element size to approximately 0.2 mr, so that a 2-inch-wide structural member occupies approximately 5.5 resolution elements, which will yield acceptable accuracy. Figure 5 is an infrared image of rib elements taken with the 10X telescope attachment from the ground. It is approximately the area indicated in the rectangle on Fig. 3. Additional magnification is possible by using the radiometer 2.1 electronic zoom. One drawback to the present system is that the extremely narrow total field of view of approximately $1.4 \times 1.8^\circ$ makes structural rib element identification extremely difficult with the infrared radiometer alone. For this reason, the present test procedure utilizes a standard video camera with a high magnification zoom lens hard-mounted to the same remotely controlled pan/tilt head as the infrared radiometer so that both systems may view approximately the same scene. Positive rib element identification is then established with the video image. Figure 6 is a photograph of this test assembly. The larger, black telescope attachment (seen on the left in Fig. 6) is the 10X infrared lens. Using this procedure, a relatively complete data set may be accumulated within approximately 15 minutes. It is felt that changes in ambient conditions and radiometer reference levels are acceptably small during this period.

Radiometer calibration is an important consideration in the present application. It is necessary to establish a known reference level against which rib radiositivities may be measured. It is also necessary to accurately determine radiosity attenuation due to such factors as telescope optical transmissivity and spectral atmospheric absorption under the local, instantaneous conditions. Initial tests of the present method, described in section V, utilized a long, thick-wall aluminum tube, closed at one end, with a 2-1/2-inch-diameter aperture at the other end. The interior was painted with a flat black paint, and the exterior was heavily insulated. It was estimated that cavity emissivity was approximately 0.98, based on Ref. 3. Two thermocouples embedded in the closed end cap indicated the temperature. This "homemade," uncontrolled reference source was placed at the base of the antenna structure and viewed through the same optics from the same distance as the antenna structural members. The test procedure called for establishment of a reference level before the actual data collection. The radiometer sensitivity was adjusted so that temperature differences of approximately 16°C on "black" objects at around room temperature without the telescope attachment would

cause the display of radiosity on the video monitor to vary from saturated black to saturated white, with 64 discrete grey levels in between. After the reference device reached a stable temperature (near ambient), a single reference level was established and recorded on videotape. The instrument was then used to measure and record on videotape radiosities of the antenna structural members. The base reference level was again checked after the data collection.

Calibration procedures in future tests will probably utilize a controlled black temperature reference source. Accuracy will be improved if a complete calibration curve, defining observed radiosity versus source temperature, is obtained under identical operating conditions. Simplifying assumptions discussed in section II will then be used to estimate the true emitted energy, and hence, actual rib temperatures.

V. Preliminary Results

An initial feasibility test was performed at the DSS 13 site on 7/30/81 utilizing the ground-based infrared inspection procedure outlined in section IV. The antenna was forced to track the anticipated sun position with a four-hour lead. The purpose was to attempt to keep approximately the same solar insolation and shadow pattern on the antenna and rib structure, while at the same time inducing a severe thermal gradient condition. As a result of the tracking offset, approximately 1/3 of the dish surface was directly illuminated, and the rest shadowed. A significant portion of the rib structure was also in the sun.

Figures 7 and 8 represent an isothermal map of DSS 13 rib structure under the test conditions indicated. Figure 7 indicates the temperatures on the rib elements directly beneath the dish surface, Fig 8 indicates lower rib elements. Only about 1/2 to 2/3 of the antenna rib structure could be measured from the single IR camera ground position. Thermal data displayed are preliminary, and have not yet been fully reduced according to the assumptions and procedures outlined in Section II. Future tests will utilize more exact data reduction techniques, and may be taken from multiple ground positions to provide more complete spatial coverage (if required).

- The entire surface of the antenna at DSS 13 is painted with a thermally diffuse white paint made by dissolving titanium

oxide, zinc oxide, and diatomaceous silica in a modified polyester. The paint is white in the visible, so that it reflects most solar energy, but has a high emissivity in the longer infrared wavelengths. The rib structure below the antenna surface is also covered with this paint. However, there is a red painted band around the extreme outer rim of the antenna, and a much higher temperature at this edge was observed due to the increased solar absorptance of the red color. An unpainted steel platform was added below the rib structure, and a 5-10°C difference between the painted and unpainted surface was observed.

Figures 7 and 8 indicate maximum rib thermal gradients on the order of 6°C under the instantaneous ambient conditions. The data will be used in a structural computer code in an attempt to calculate thermally induced antenna surface deflections and resulting RF gain losses. Additional directions of the thermal test program are discussed in section VI.

VI. Conclusions and Future Work

Experience gathered to date with the infrared camera indicates that the radiometric technique may be used to provide accurate thermal gradient data about the large and complex antenna structures. Additional refinements to the overall test procedure are expected to further reduce the effort and cost of future analyses.

Current plans call for additional thermal analyses of the 32-m-diameter Venus antenna (DSS 13) while employing several different antenna solar tracking modes in the attempt to determine worst possible case thermal gradients. The relative speed of the method should enable thermal transient investigation over the course of a day. It is also planned to study the effects on thermal gradients under several environmental conditions throughout the year. Where possible, other antennas of varying sizes within the DSCC will be studied experimentally, and attempts will be made to correlate antenna size/complexity with observed thermal gradients. The study may also be extended to cover other portions of the antennas, such as the subreflector, quadrupod legs, etc. Structural/thermal analysis interface requirements will be defined to permit convenient use of thermal gradient data in a timely manner.

References

- 1 Siegel, R , and Howell, J , *Thermal Radiation Heat Transfer*, McGraw-Hill Book Company, 1972
2. Lloyd, J M , *Thermal Imaging Systems*, Plenum Press, 1975
- 3 Wolfe, W L., *The Infrared Handbook*, prepared for the Infrared Information and Analysis Center (IRIA), Environmental Research Institute of Michigan
4. Ware, J C , “Clear Sky Temperature,” presented at the ISES Meeting, Fort Collins, Colorado, August 1974.

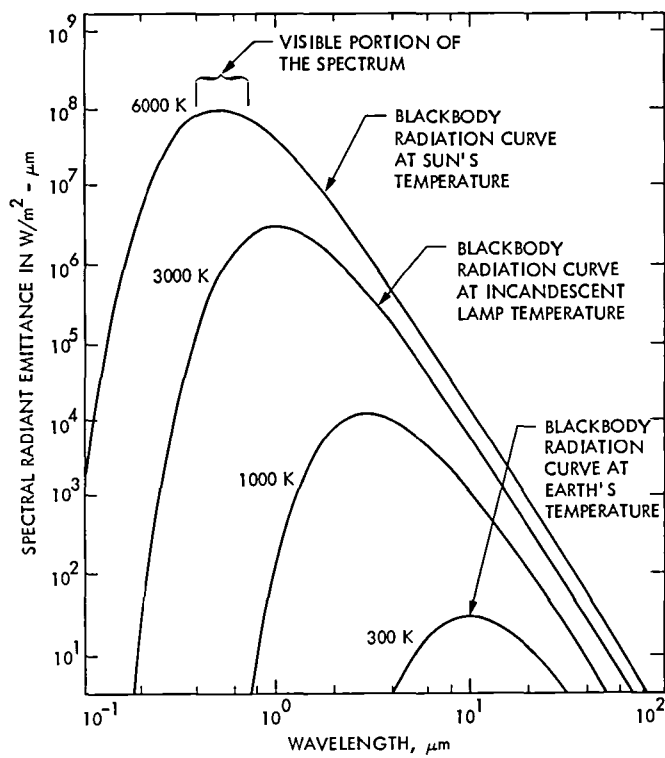


Fig 1. Spectral distribution of energy radiated from a black body at various temperatures

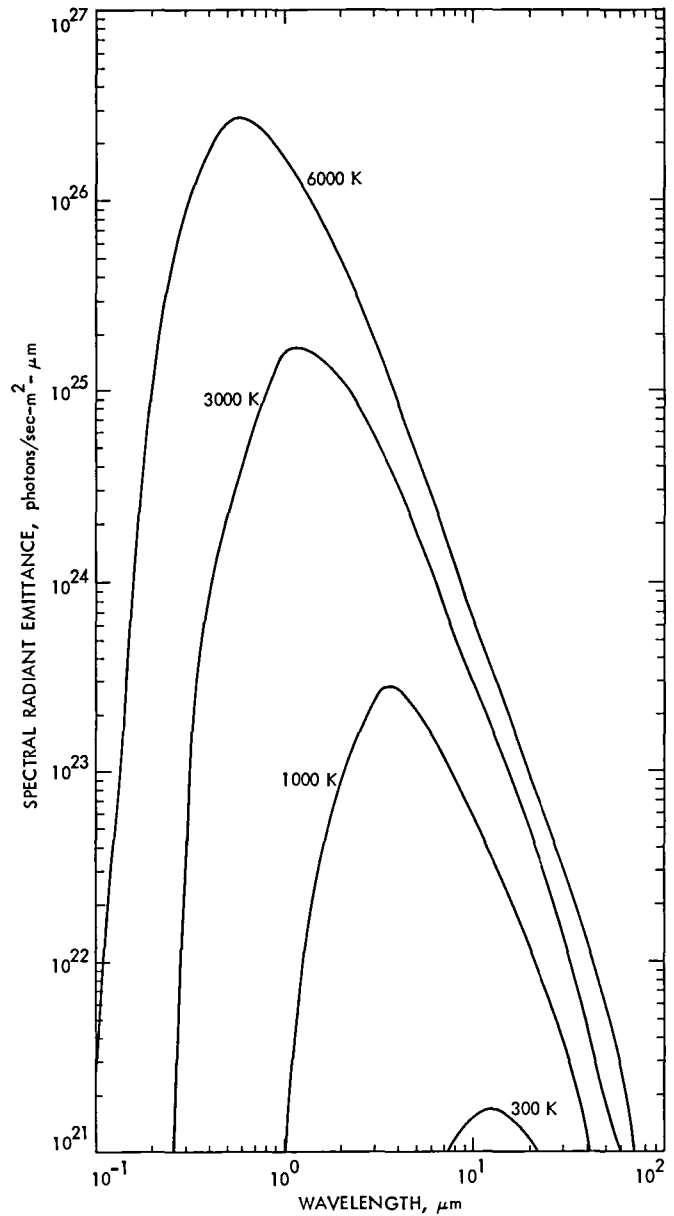


Fig. 2. Spectral distribution of photons radiated from a black body at various temperatures

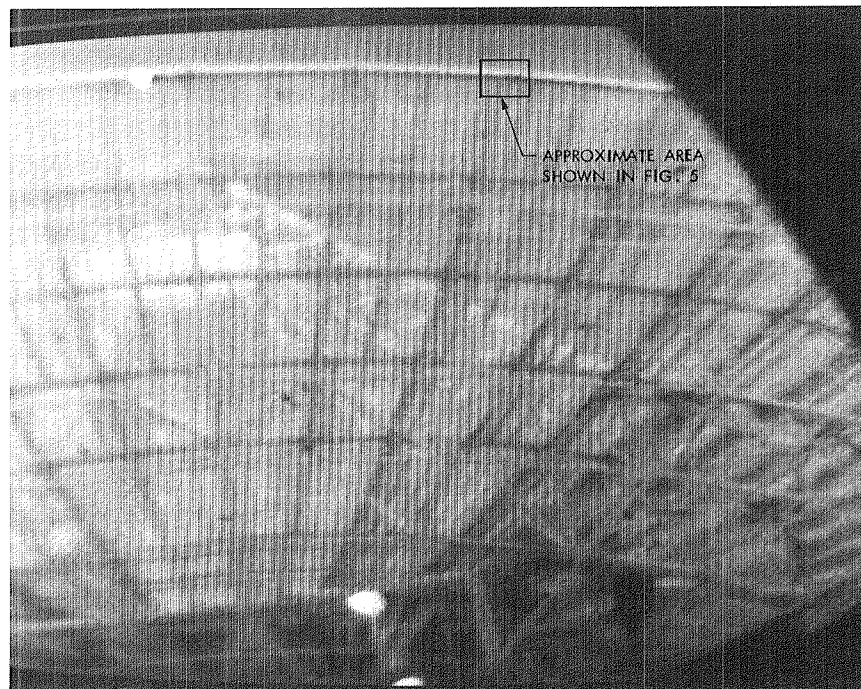


Fig. 3. Video picture of DSS 13



Fig. 4. Closeup infrared image of rib joint from cherry picker bucket, using iso-radiosity enhancement mode

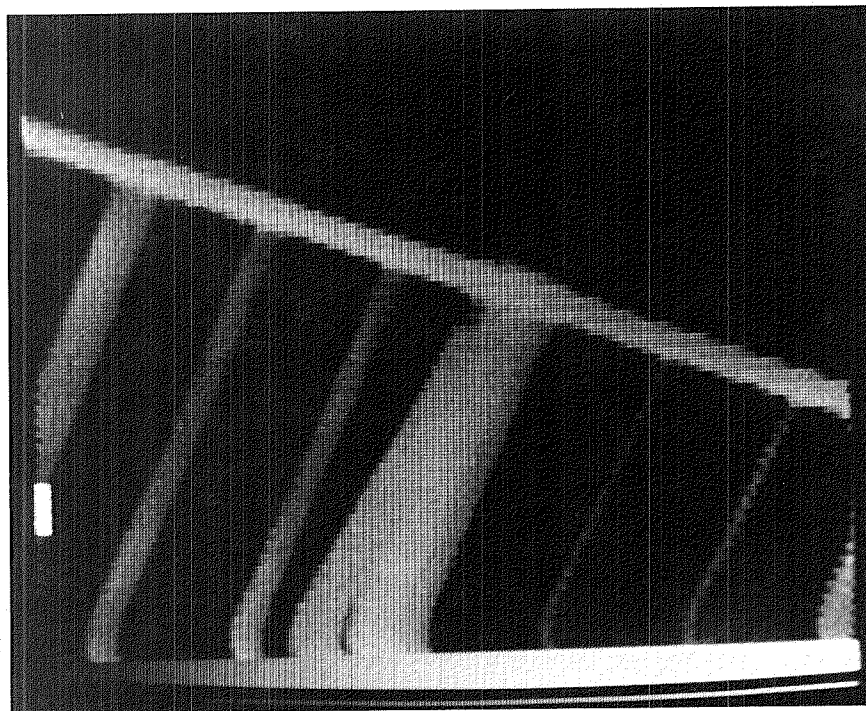


Fig 5. Infrared image of rib joint using ground-based 10× telescope from approximately 150-foot range

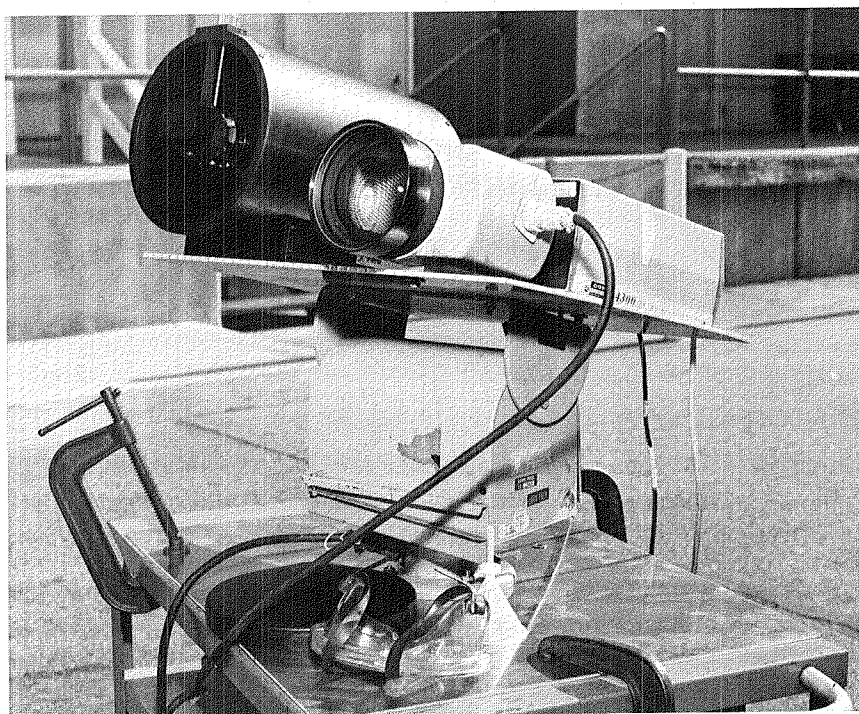


Fig. 6. Test assembly

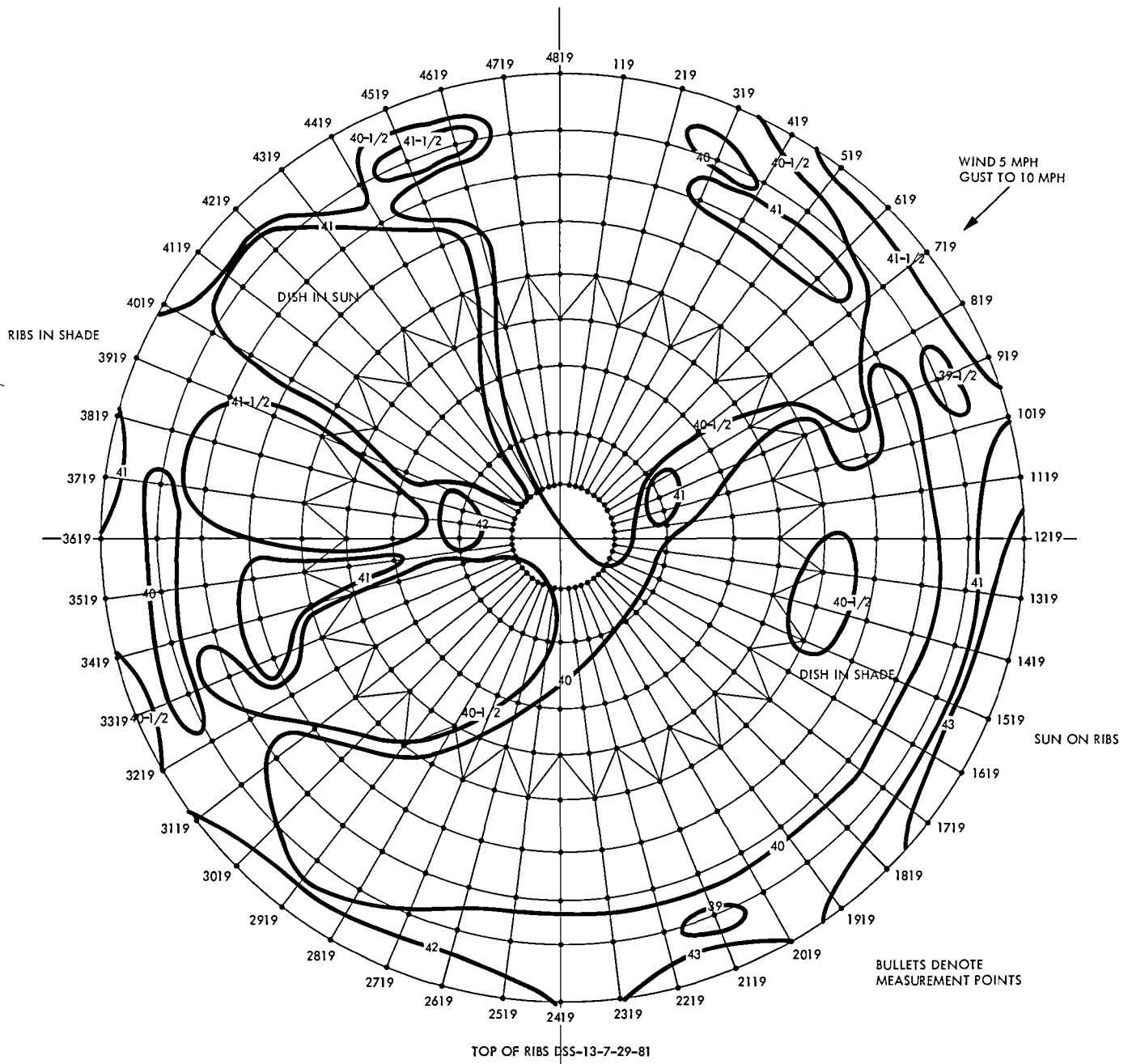


Fig 7. Isothermal rib temperature map of DSS 13, for rib joints immediately beneath dish surface

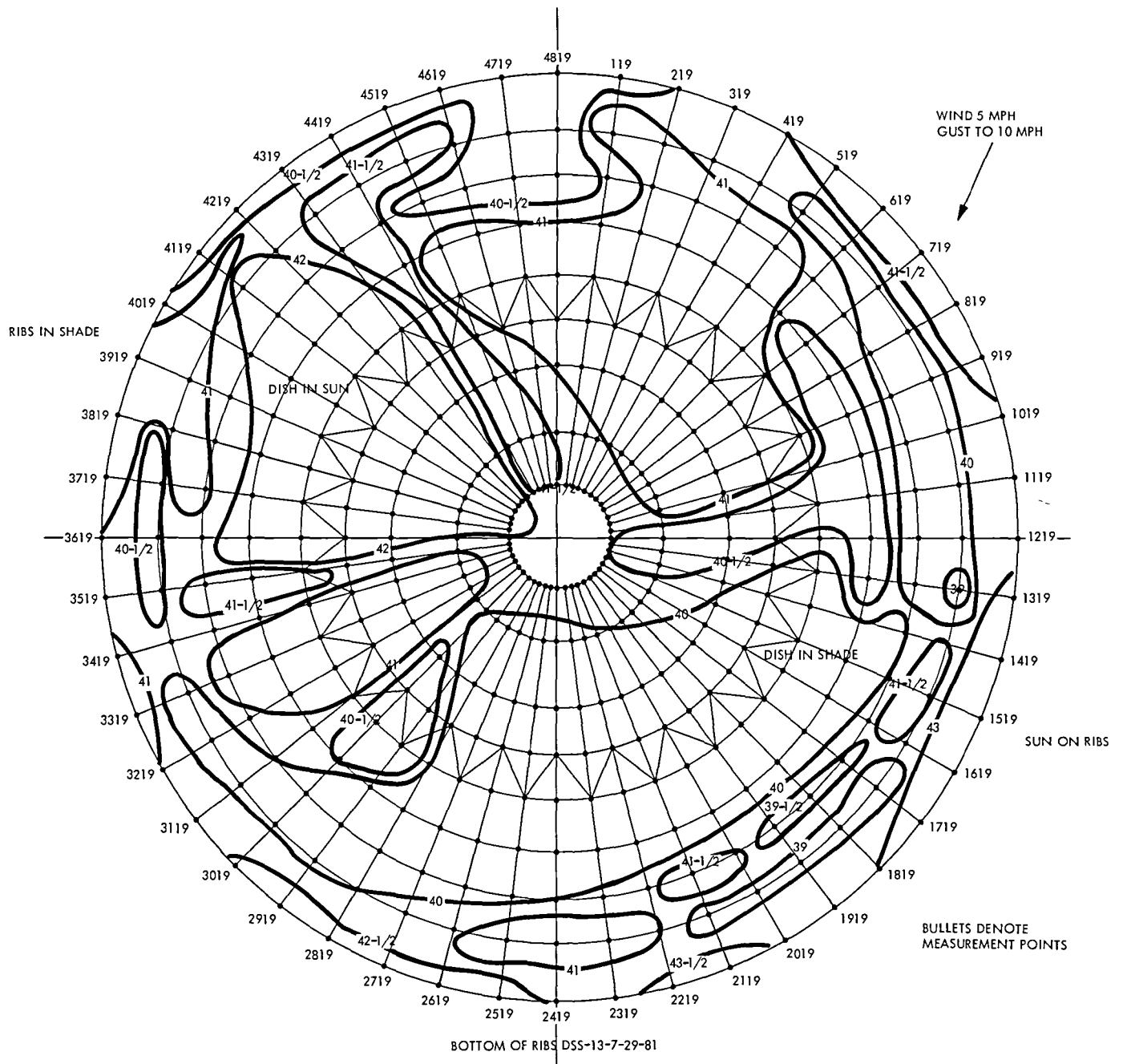


Fig. 8. Isothermal rib temperature map of DSS 13, for lower rib joints

X-Band Narrow-Beam Radiometer for DSS 13

M M Franco and S D Slobin
Radio Frequency and Microwave Subsystems Section

C T Stelzried
TDA Mission Support

A narrow-beam X-band (8 448 GHz) radiometer has been designed, constructed, and tested at JPL. This radiometer will be operated at the Goldstone Deep Space Station (DSS 13). Continuous calibration is obtained by switching the receiver sequentially among the hot termination, ambient termination and antenna positions. The measured antenna noise temperatures (10-second total integration time) have a standard deviation of about 0.4 K, which is in good agreement with theory. It is felt that radiometer performance is satisfactory for the purpose of making atmospheric noise temperature measurements.

I. Introduction

A narrow-beam X-band (8 448 GHz) radiometer has been designed, constructed, and tested at JPL. This radiometer will be installed and operated at the Goldstone Venus Deep Space Station (DSS 13) for the purpose of making atmospheric microwave noise temperature measurements. The radiometer input is automatically switched among a hot termination, ambient termination, and the antenna. This system will augment the wide-beam X-band measurements of atmospheric noise temperature also made at DSS 13 (Refs 1 and 2). These measurements are used to generate statistics of noise temperature increase above baseline values for use in Deep Space Network communication-link design tables. This report describes the radiometer and gives an example of performance based on preliminary JPL test data.

II. Radiometer Description

The antenna used for the narrow-beam system is a 6-foot-diameter (1.8-meter) dish with a focal-point feed (Fig 1).

This antenna will have a beamwidth of 1.4 degrees, the existing wide-beam system has a beamwidth of 15 degrees. A metal shroud, lined with absorbing material, surrounds the aperture of the dish. This addition is purported to give both low side-lobes and to minimize any changes in system noise temperature due to scattering and spillover effects. Antennas of the type described here are used commercially in horizontal line-of-sight applications, and the shroud may or may not be an advantage in the DSS 13 application where the antenna nearly always points upward at an elevation angle of 30 degrees or more. Antenna pattern and noise temperature tests will be made both with and without the shroud installed on the antenna.

The radiometer electronics package is mounted on the back of the antenna structure (Fig 2), which itself is mounted on an elevation over azimuth positioner. Figure 3 shows the block diagram of the radiometer. This radiometer is continuously calibrated by switching the receiver sequentially among the hot termination, ambient termination, and antenna positions. The receiver consists of an 8 448 GHz mixer and a 10-100

MHz IF amplifier with appropriate gain to interface with the square-law detector (Refs 3-5)

III. Radiometer Calibration

The programmer (Fig 3) controls the switching sequence and continuously calculates the antenna noise temperature. This calculation utilizes the output of the square-law detector and the appropriate termination physical temperatures. Antenna noise temperature is computed from

$$T_A = T_{AMB} - \left(\frac{T_{HOT} - T_{AMB}}{V_{HOT} - V_{AMB}} \right) (V_{AMB} - V_{ANT}) \quad (1)$$

where

T_{AMB} = radiometric noise temperature of ambient termination, K

T_{HOT} = radiometric noise temperature of hot termination, K

V_{HOT} = detector output when switched to the hot termination, V

V_{AMB} = detector output when switched to the ambient termination, V

V_{ANT} = detector output when switched to the antenna, V

The radiometer gain, $(T_{HOT} - T_{AMB})/(V_{HOT} - V_{AMB})$ K/V, is continuously monitored to assess radiometer performance. The physical temperature of the hot termination is not generally its radiometric noise temperature. The radiometric noise temperature of the ambient termination is equal to its physical temperature if all lossy elements of the termination and associated waveguide are at the same physical temperature. The correction for the hot termination is found by means of tipping curve measurements. T_{HOT} in Eq (1) is adjusted to correct the antenna temperatures as determined at various elevation angles.

IV. Radiometer Analysis

The resolution (1σ) of a radiometer measurement is obtained from Eq (1)

where

$\sigma_{T_{A,TP}}$ = total power radiometer resolution when switched to the antenna, K

$$= \frac{T_{op,A}}{\sqrt{B\tau}}$$

$\sigma_{T_{HOT,TP}}$ = total power radiometer resolution when switched to the hot termination, K

$$= \frac{T_{op,HOT}}{\sqrt{B\tau}}$$

$\sigma_{T_{AMB,TP}}$ = total power radiometer resolution when switched to the ambient termination, K

$$= \frac{T_{op,AMB}}{\sqrt{B\tau}}$$

T_{ANT} = antenna noise temperature, K,

and

T_{op} = total system noise temperature on any signal source, K

B = IF bandwidth, Hz

τ = measurement integration time, sec

Substituting the above (and using $T_{ANT} = 10$ K, $T_{AMB} = 310$ K, $T_{HOT} = 425$ K, $\tau \cong 10$ sec, and $B = 10$ MHz) into Eq. (2) results in a theoretical antenna noise temperature measurement resolution (σ_{T_A}) of approximately 0.4 K.

V. Radiometer Performance

Atmospheric noise temperature tipping curves have been performed during radiometer tests at JPL. The data taken on September 21, 1981 are reported here. The sky was clear, the ground relative humidity was 55% and the ground temperature was 32°C. The tipping curves were performed by moving the antenna between zenith and 30° elevation.

The radiometer was switched to each termination position for approximately 10 secs. The calibration sequence takes $\cong 40$ secs. These data are shown in Fig 4.

$$\sigma_{T_A} = \sqrt{\left(\sigma_{T_{A,TP}} \right)^2 + \left(\frac{T_{AMB} - T_{ANT}}{T_{HOT} - T_{AMB}} \right)^2 \left(\sigma_{T_{HOT,TP}} \right)^2 + \left(\frac{T_{HOT} - T_{ANT}}{T_{HOT} - T_{AMB}} \right)^2 \left(\sigma_{T_{AMB,TP}} \right)^2}$$

(2)

The standard deviation of the antenna noise temperature measurements at zenith is approximately 0.4 K, which agrees closely with theory. The separately averaged temperatures (≈ 20 min) at zenith agree with each other within 0.2 K, which is reasonable for this data since the atmosphere itself may have changed during the two-hour test period. It is felt that radiometer performance is satisfactory.

An analysis of Fig. 4 indicates several calibration problems which need to be solved. First, the zenith antenna temperature values average around 1.6 kelvins. This is clearly in error, since the cosmic background contribution is 2.7 kelvins and a generally accepted clear sky atmospheric contribution is about 2.5 to 2.8 kelvins, for a total of approximately 5.5 kelvins.

Ground and waveguide contributions may add another 5 kelvins for a minimum total expected antenna temperature of about 10 kelvins. This error is clearly due to the use of the hot termination physical temperature instead of a calibrated radiometric noise temperature, and an unmodelled ground and waveguide contribution. A clue to ground contribution is seen from the approximately 3.3-kelvin differences between zenith and 30-degree noise temperatures. Since the atmospheric path length increases by one atmosphere (from 1 to 2 atmospheres) in the tipping curve method, a 2.5- to 2.8-kelvin noise temperature would be expected. The excess noise temperature is probably due to increasing ground contribution and will be modelled to obtain a total radiometer calibration. The corrections will be incorporated into Eq. (1).

Acknowledgment

The authors wish to thank Kenneth B. Wallace, now with TRW, Redondo Beach, California, for carrying out much of the initial work on the narrow-beam radiometer.

References

- 1 Reid, M. S., Parham, O. B., and Gardner, R. A., "An X-Band Radiometer for the Microwave Weather Project," *DSN Progress Report 42-29*, pp. 54-59, Jet Propulsion Laboratory, Pasadena, Calif.
- 2 Slobin, S. D., Stelzried, C. T., Andres, E. M., and Franco, M. M., "X-Band Atmospheric Noise Temperature Statistics at Goldstone, DSS 13, 1979 and 1980, and Clear Air Noise Temperature Models for Goldstone," *TDA Progress Report 42-64*, pp. 161-167, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1981.
- 3 Reid, M. S., Gardner, R. A., and Stelzried, C. T., "A New Broadband Square Law Detector," Technical Report 32-1526, Vol. XVI, pp. 78-86, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1973.
- 4 Reid, M. S., Gardner, R. A., and Stelzried, C. T., "Improvement in the Accuracy of the New Broadband Square Law Detector," Technical Report 32-1526, Vol. XVIII, pp. 94-98, Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1973.
- 5 Gardner, R. A., Stelzried, C. T., and Reid, M. S., "Radiometric Applications of the New Broadband Square Law Detector," Technical Report 32-1526, Vol. XIX, pp. 89-92, Jet Propulsion Laboratory, Pasadena, Calif., February 15, 1974.

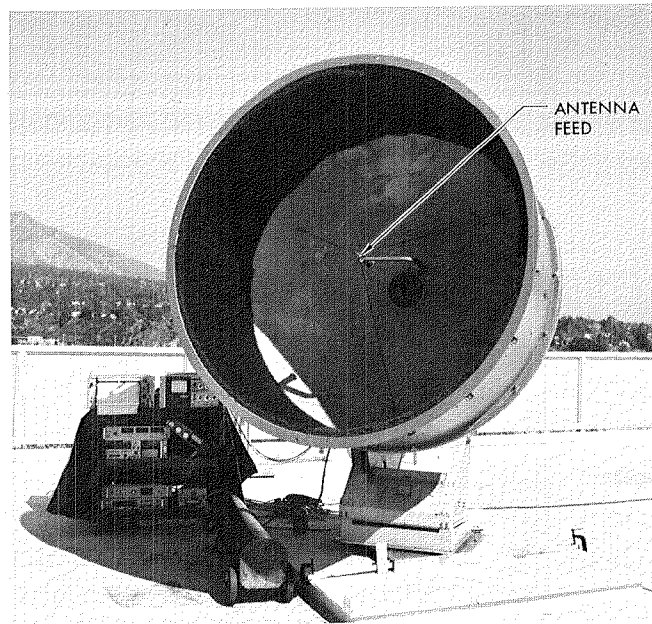


Fig. 1. Narrow-beam X-band radiometer (front view)

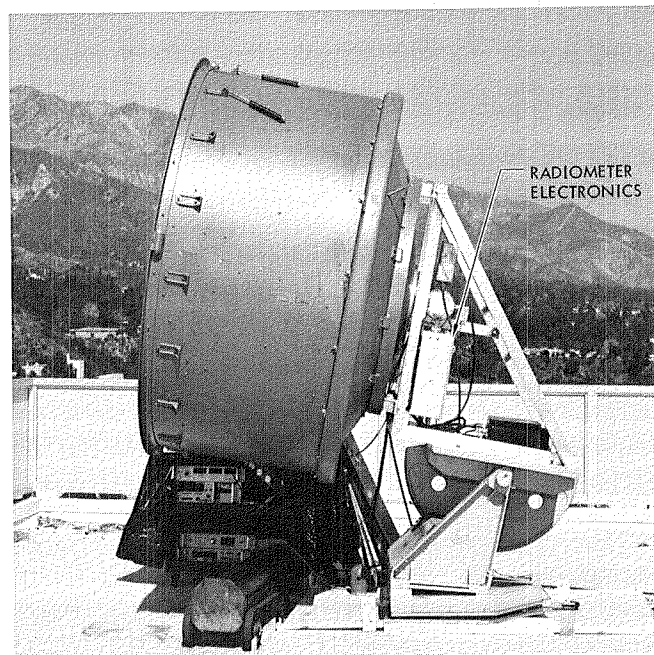


Fig. 2. Narrow-beam X-band radiometer (side view)

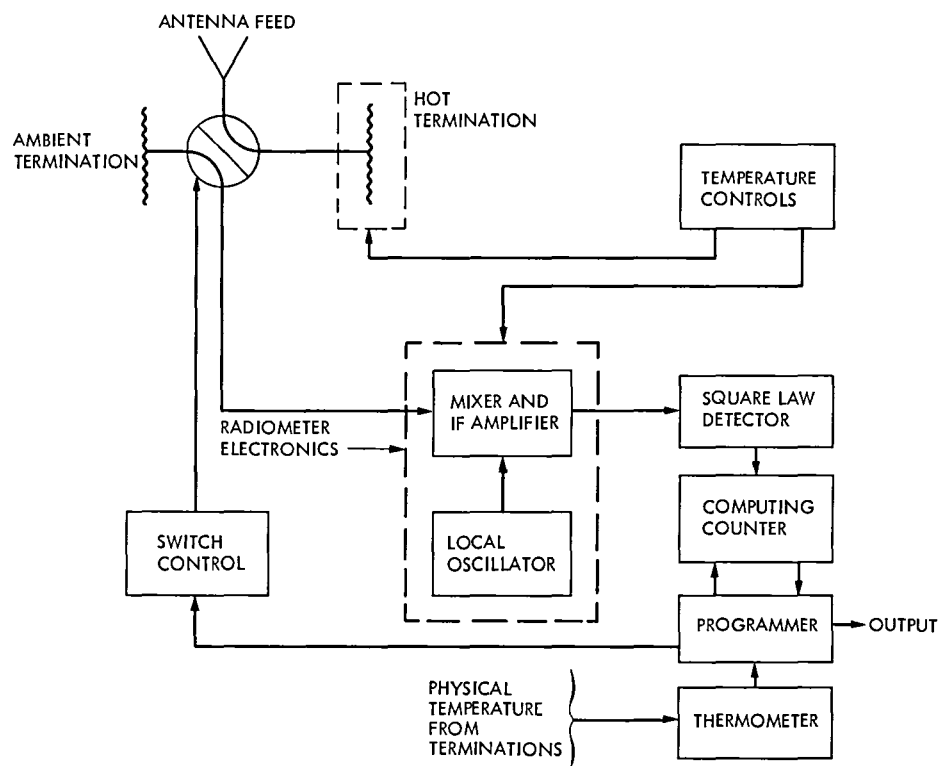
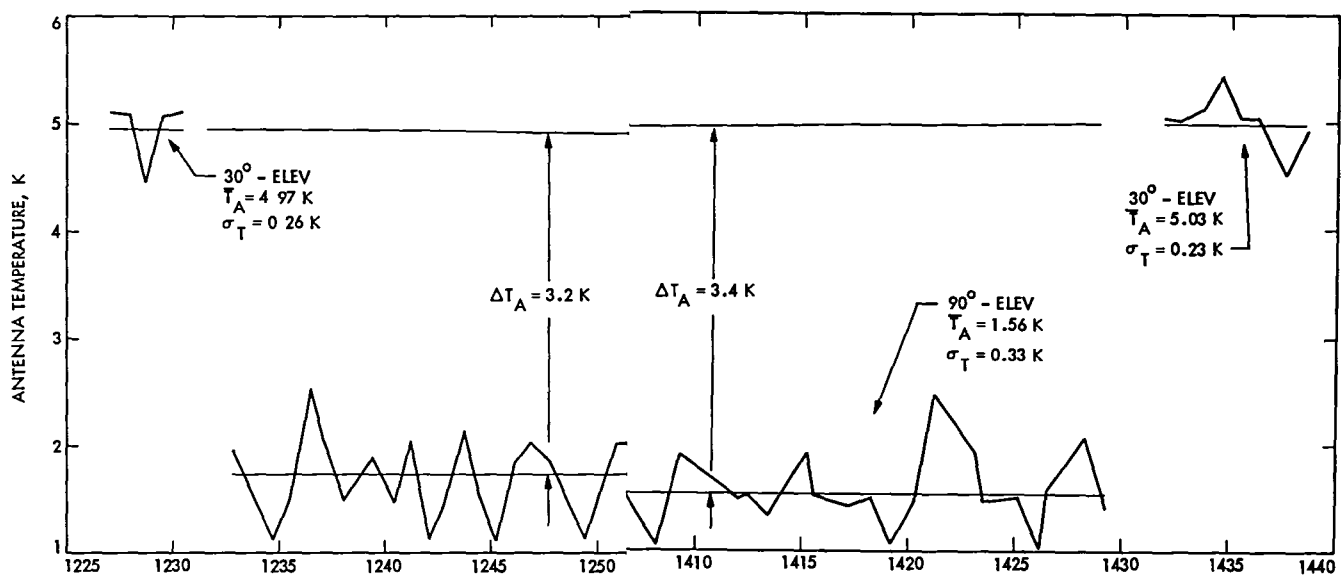


Fig. 3. Block diagram of the narrow-beam X-band radiometer



med at JPL, September 21, 1981 with the narrow-beam X-band radiometer

This Page Intentionally Left Blank

Design of a 1.5-m, 32-GHz, Clear Aperture Antenna

A Cha

Radio Frequency and Microwave Subsystems Section

Present DSN ground station antennas are of the symmetric dual-reflector type. In the last few years, an investigation has been made of alternative ground station antenna designs which have a clear aperture (no subreflector or strut blockage) and shaped reflector surfaces. In FY82, a 1.5-m, 32-GHz clear aperture antenna model will be built for experimental studies. This article describes the underlying considerations leading to the determination of the parameters defining the model antenna geometry. Detailed analysis of the model electrical characteristics will be reported at a later time.

I. Introduction

Present DSN ground station antennas are of the symmetric Cassegrainian dual reflector type. In the last few years, an investigation has been made of alternative ground station antenna designs which have a clear aperture (no subreflector nor strut blockage) and shaped reflector surfaces. The new configurations are shown to have a potential 2 to 3 dB G/T (gain/noise temperature ratio) performance enhancement compared to the present symmetric reflector stations (Ref 1). Figure 1 affords a schematic illustration of the new antenna under study along with a conventional Cassegrain.

Significant advances in the synthesis, analysis and performance optimization of this new type of antenna have been accomplished since the investigation was initiated in 1978. Parallel studies on the structural design of asymmetrical reflectors were undertaken in the same period. With this much theoretical development, a new task to design, fabricate and test a 1.5-m clear aperture antenna model was started in FY81. The goal is to provide experimental verification of the predicted superior performance of the new antenna design.

This article describes the underlying considerations leading to the determination of the parameters defining the antenna geometry. We will start by describing briefly the reflector synthesis and analysis techniques used in the design process.

II. Optical Synthesis of the Offset Dual Shaped Reflector—the First Design Iteration

An approximate synthesis of the offset dual-shaped reflector problem based on ray optics laws was developed in 1979 (Refs 2, 3). The geometry of the problem is shown in Fig 2. In an exact synthesis, main and subreflector surfaces which convert a spherical wave emanating from the feed to a planar wave with prescribed amplitude and phase distribution in the main reflector aperture are generated. In our approximate approach, the output reflector aperture amplitude distribution has a small deviation from the prescribed distribution although the aperture phase distribution always follows the prescribed one exactly (generally a uniform phase distribution is desired).

There are a number of input parameters in the synthesis theory, Z_0 , R_0 , ρ_M , A , Ω and θ_M , as shown in Fig 2. These parameters provide the flexibility to determine a preliminary configuration with desirable main-to-subreflector-diameter ratio and relative dislocations between the feed, the subreflector and the main reflector. In this iteration, additional input parameters p , ρ_{MM} , n (see following) are held to constant values.

The parameter n is used in the synthesis to specify a tapered input feed illumination of the form

$$I(\theta) = \cos^{n-1} \theta \quad (1)$$

In practice, the value $n = 85$ at which $I(\theta)$ approximates the pattern of a corrugated horn best is used in the first iteration.

A particularly interesting form of output aperture energy distribution is

$$V(\rho) = \left(1 - \frac{\rho^2}{\rho_{MM}^2}\right)^p \quad (2)$$

where $\rho_{MM} > \rho_M$ and $0 \leq p \leq 1$. The parameter $\rho_{MM} > \rho_M$ is introduced to avoid a mathematical singularity which would arise in the synthesis theory if ρ_{MM} in Eq (2) is replaced by ρ_M . It has been found from experience that ρ_{MM} should be set to a fixed, nominally larger value than ρ_M . The value of the exponent p is arbitrarily fixed for the first design iteration, e.g., $p = 0$. Its value will be determined in a second iteration which enables one to select the optimal design based on the more accurate diffraction analysis (as opposed to optics). It is significant to note that changing the p value within its range 0 to 1 does not usually change the reflector geometry and surface profiles appreciably to necessitate new backup structure design. At the conclusion of the first design iteration, the basic geometry and dimensions of the antenna can be considered fixed. The design of the backup structure can be initiated concurrently with the second design iteration (diffraction design). Figure 3 shows a profile of the 1.5-m model obtained from this design cycle which was the basis for the initial mechanical and structural design.

III. Diffraction Analysis and Efficiency Computations—the Second Design Iteration

In this second iteration, we will analyze the diffraction characteristics and compute the efficiency of reflector designs corresponding to different p values in Eq (2), holding the other reflector input parameters Z_0 , R_0 , etc. constant. Fig-

ure 4 shows an overview of this iteration cycle. Each efficiency term represents a loss mechanism within the antenna system. The overall antenna efficiency η_T is

$$\eta_T = \eta_{FS} \eta_{RS} \eta_P \eta_X \eta_I \eta_{RMS} \eta_C \quad (3)$$

The terms η_{RMS} and η_C are losses due to surface RMS error and dissipation in the antenna and feed. These are not included in the diffraction and efficiency analysis. In addition, one can obtain a first-order estimate of the antenna noise temperature at zenith from the rear spillover efficiency as

$$T_a = (1 - \eta_{RS}) \times 240 \text{ kelvins} \quad (4)$$

At the end of this iteration, we will have determined the final reflector profiles of an optimal gain or G/T design, the antenna efficiency (gain) and antenna noise temperature.

Note that Eq (2) describes a nearly uniform aperture energy distribution with a taper near the aperture edge. It is not altogether apparent how an optimal gain or G/T design can result from specifying the particular aperture illumination function. In Refs 4, 5, and 6 it is shown that this design strategy allows one to examine the tradeoff between a worsening illumination efficiency and an improving rear diffraction spillover efficiency as the illumination taper is increased (p increased) and arrive at an optimal design in gain or G/T.

IV. A Summary of Design Guidelines

For convenient reference, we provide the list below:

- (1) The feed is a corrugated horn with 22 dB gain.
- (2) The main and subreflector diameters are 1.5 and 0.45 m respectively.
- (3) The input parameters to the reflector synthesis program are $n = 85$, $p = 0.5$, $R_0 = 3.0$, $Z_0 = 0$, $A = 4.5$, $\rho_M = 3.0$, $\rho_{MM} = 3.000001$, $\theta_M = 16^\circ$, $\Omega = 0^\circ$.
- (4) Model test frequency is 32 GHz.
- (5) Reflector RMS surface tolerance is 0.2 mm.
- (6) Alignment accuracy between feed, subreflector and main reflector is 0.4 mm.

V. Feed for the Model

The feed design selected is that of a corrugated horn with an on-axis gain of 22 dB. The n -value at which $\cos^n \theta$ best approximates the corrugated horn power pattern in the least RMS error sense is found to be 85. A comparison of these two patterns is shown in Fig 5a. The subreflector illumination

angle θ_M subtended at the feed is chosen to be 16° . At this angle, 96% of the feed radiated energy is intercepted by the subreflector; i.e., $\eta_{FS} = 0.96$. Although η_{FS} can be further improved by using a larger θ_M , this is attenuated by an expected lower phase efficiency η_p due to worsening variation in the feed phase pattern at larger angles (Fig. 5b). An iteration cycle to determine the optimal feed illumination angle θ_M and modification of the synthesis theory to incorporate arbitrary feed illumination pattern (instead of $\cos^n\theta$) were expected to yield at best token improvement and were not undertaken.

VI. The Reflector Design

The size of the subreflector diameter is the starting point of the reflector design. At the planned test frequency of 32 GHz, a 0.45-m subreflector is close to 50 wavelengths. This is considered to be of sufficient size for the optical laws, which form the foundation of the reflector synthesis theory, to be good approximations to diffraction theories. The main reflector diameter is 1.5 m, which corresponds to a main-to-subreflector-diameter ratio of 3.3 to 1. This ratio is lower than what one would normally use in a two-reflector system but is selected to hold down the fabrication cost. Note the input parameters R_0 , Z_0 , A , ρ_M and ρ_{MM} to the computer program, as given in Section IV, are dimensionless. These and the reflector surface coordinates generated by the program are scaled according to the reflector size. For example, in the case of the model design, the scale factor is 0.5 m, which brings $\rho_M = 3.0$ to $\rho_M = 1.5$ m.

The parameters Z_0 , R_0 , A , Ω , θ_M largely determine the relative dispositions of the feed, subreflector and main reflector. Care should be taken to ensure that a structurally acceptable and totally blockage-free package results from the choice of these. The feed tilt angle Ω is considered a less critical parameter in determining the model performance characteristics relevant to DSN operations and is set to 0° . Some evidences in the literature (Refs. 7, 8) suggest that this tilt angle may have a profound bearing on the polarization characteristics of offset dual reflectors. If the intended

usage of the antenna requires high polarization purity as in some earth orbit satellite ground stations, then an additional design iteration cycle determining the optimal feed tilt angle would be in order. In the DSN mode of operations, however, the primary figure of merit G/T is largely unaffected by the selection of this parameter, as the main effect of depolarization is a slight degradation in gain, on the order of 1% or less. This observation is supported by the diffraction and efficiency analysis which will be reported separately. With values of θ_M and Ω thus fixed, the first iteration involves varying input parameters Z_0 , R_0 and A only.

The parameter p , the exponent in Eq. (2), is determined from diffraction and efficiency analysis in the second design iteration to be 0.5, based on the optimal gain criterion. The computed aperture (or area) efficiency of the model is 86% exclusive of η_{RMS} and η_C .

The specified RMS surface tolerance of 0.2 mm is roughly one fiftieth of a wavelength at 32 GHz. It is believed that this tolerance and the specified alignment accuracy would limit the gain loss due to mechanical and structural imperfections to 2% or less. No firm guides can be found in the literature on how tight these tolerances should be. We are providing precision micrometer-type mounts for both the feed and the subreflector. These adjustments will allow us to peak up the antenna gain during tests and to study how antenna characteristics are affected by misalignment. Figure 6 shows the 1.5-m clear aperture antenna with its backup structure and a base plate. The whole assembly can be mounted on a standard antenna range pedestal for testing.

VII. Summary

The design of the 1.5-m clear aperture antenna has been presented. The procedure consists of two iterative cycles. In the first iteration, the geometry of the reflector is largely determined based on optics laws. In the second iteration, the optimal design is determined from the more accurate diffraction analysis. Detailed performance analysis will be reported in a future article.

References

1. Cha, A. G., Bathker, D., and Williams, J., "Advanced Design Concepts of Ground Station Antennas," Proceedings of 9th European Microwave Conference, Brighton, England, Sept. 1979. Also in *Advanced Antenna Technology*, edited by P. J. B. Claricoats, Microwave Exhibitions and Publishers Limited, 1981.
2. Galindo-Israel, V., Mittra, R., and Cha, A., "Aperture Amplitude and Phase Control of Offset Dual Reflectors," *IEEE Trans. Ant. Prop.*, March 1979.
3. Cha, A., Galindo-Israel, V., and Mittra, R., "Efficient Design of Offset Dual Shaped Reflectors for Antenna and Beam Waveguide Applications," AP-S International Symposium, Seattle, June, 1979.
4. Cha, A. G., Galindo-Israel, V., and Bathker, D. A., "Low Noise Dual Shaped Reflector Synthesis," AP-S International Symposium, Quebec, Canada, June 1980.
5. Cha, A. G., "Improved Designs of High Gain Low Noise Dual Shaped Reflectors," 11th European Microwave Conference, Amsterdam, Sept. 1981.
6. Cha, A. G., "Wide-Band Diffraction Improved Dual Shaped Reflectors," to appear in *IEEE Trans. Ant. Prop.*, March 1982.
7. Mizugutch, Y., Akagawa, M., and Yokoi, H., "Offset Dual Reflector Antenna," IEEE AP-S International Symposium, Amherst, Mass., 1976.
8. Albertsen, N. C., "Offset Cassegrain Antenna Optimized for Frequency Reuse," IEEE AP-S International Symposium, Stanford, Calif., 1977.

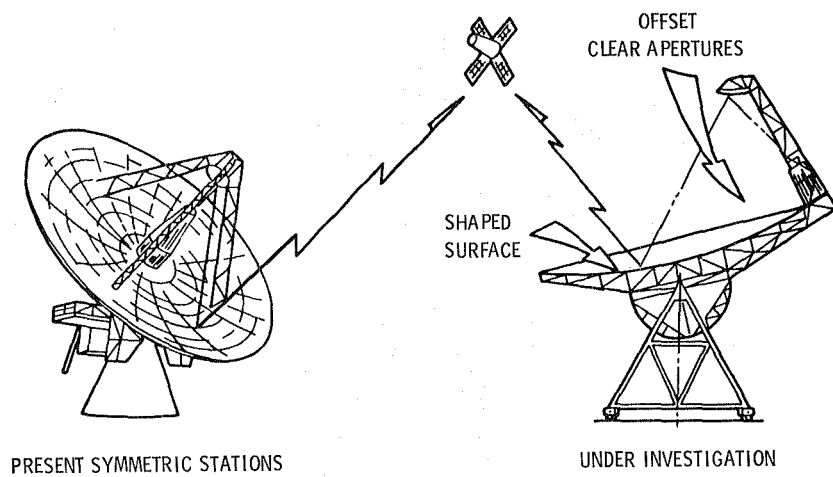


Fig. 1. Alternative reflector configurations for very low noise and high gain ground stations

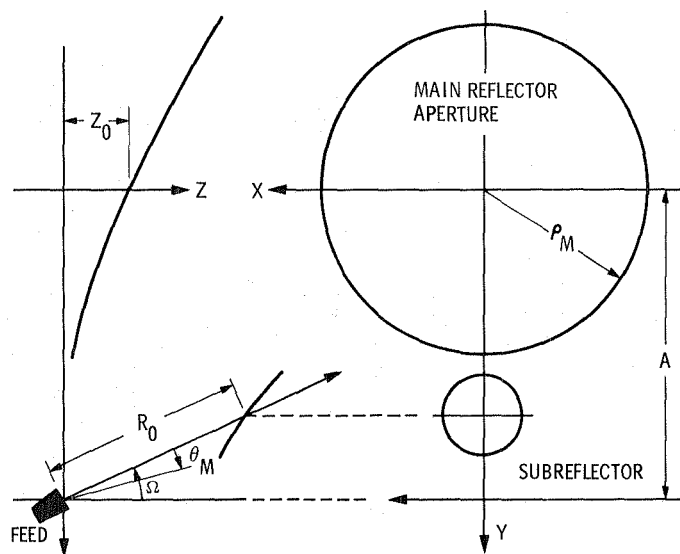


Fig. 2. Geometry of offset dual reflectors

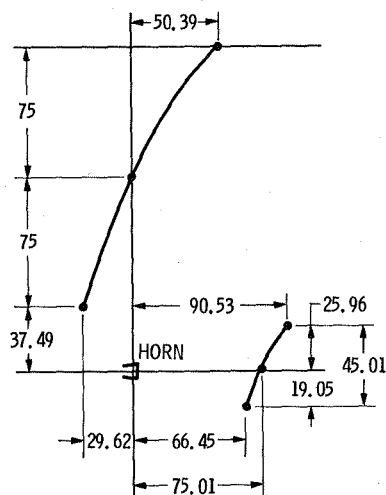


Fig. 3. Major dimensions (in cm) of the 1.5-m model

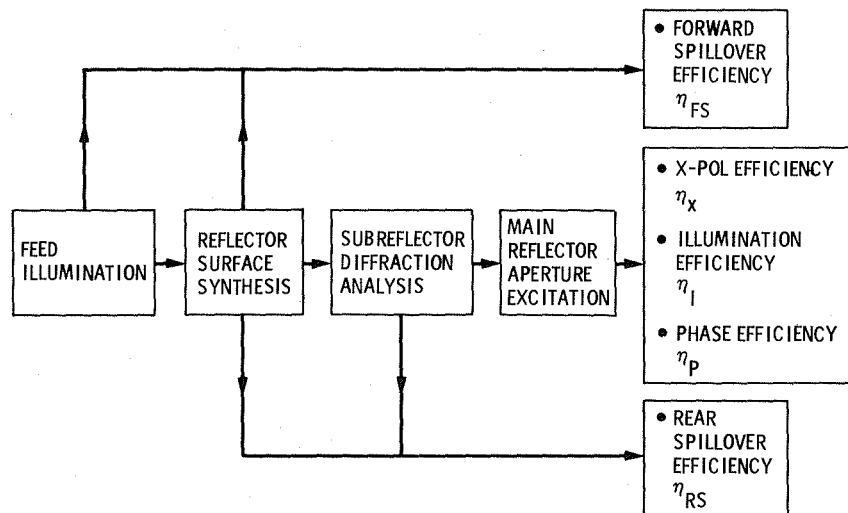


Fig. 4. Overview of reflector RF performance analysis

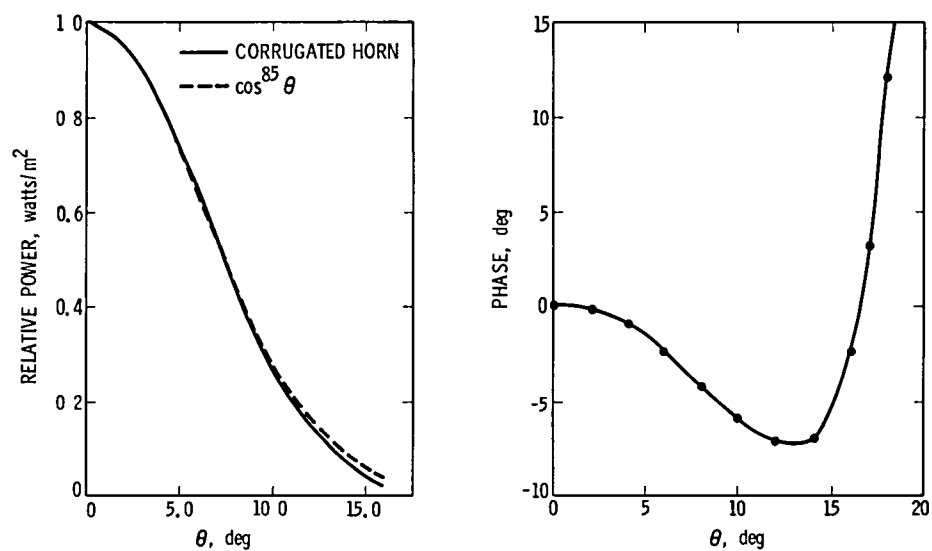


Fig. 5. Corrugated horn pattern fit by $\cos^p \theta$, (a) power patterns (b) E- and H-plane average phase of horn

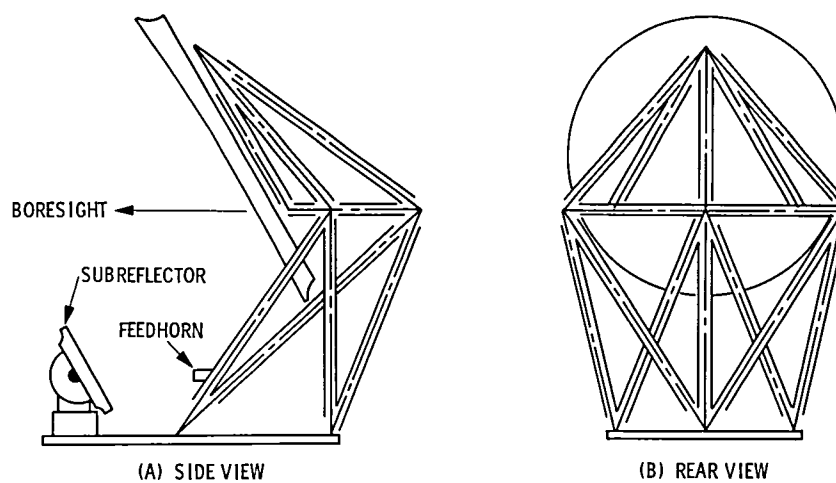


Fig. 6. 1.5-m clear aperture model and backup structure

Power Efficiency of Semiconductor Injection Lasers

J. Katz

Communications Systems Research Section

One of the advantages of semiconductor injection lasers is their potentially high power conversion efficiency. A device model and analysis of the power efficiency is presented and its dependence on the various device parameters such as length, mirror reflectivity, absorption coefficient and material conductivity is explained. Although the actual values of parameters used pertain to lasers fabricated from the AlGaAs system, the theoretical analysis is also relevant to semiconductor injection lasers of other atomic systems.

I. Introduction

Optimization of the overall efficiency of injection laser diodes is very important, especially for reliable long life optical space communication systems (Refs 1, 2). This problem has been addressed by several authors, but their analyses contained various approximations and assumptions which are not always valid. In the treatment of Sommers (Ref 3), the use of non-local rate equations is implied. This renders his results inaccurate in the region of low mirror reflectivities, which is the region where the efficiency is maximized in many cases. Whiteaway and Thompson (Ref 4) take into account the spatial distribution of the light intensity along the laser cavity (Ref 5). However, they assume that the laser operates very high above threshold, thus neglecting the spontaneous recombination of carriers. It can be shown that under some relevant physical conditions, the point of maximum efficiency is obtained when the laser is operated at current levels which are only a few times the threshold value.

This article treats the problem of efficiency optimization without the above assumptions except for neglecting the coupling of spontaneous emission into the lasing mode. This term can clearly be neglected in lasers not intended to operate at near threshold power levels. Other assumptions — also

used in other works — are that the current injection is uniform along the laser and that thermal effects can be neglected because of low duty cycle operation.

In addition, several important constraints are included. From the results of previous work done on catastrophic mirror degradation in laser diodes (Refs 6, 7), it is likely that the limiting parameter is the total photon density in the vicinity of the mirrors and not just the output light intensity. The effects of this constraint are described herein. There might also be other constraints on the laser operation besides that of the optical power. For example, the maximum current density can be a constraint, due to either thermal or reliability (Ref 8) considerations. Another possibility is a constraint on the maximum available current from the primary power source, which can be imposed by overall system specifications. The effects of these constraints will also be considered.

Sections II and III of this report outline the derivation of the expressions for the output power and the power efficiency, respectively, of the laser diode. Section IV presents results of some numerical examples and a discussion. The results show that the different physical and geometrical parameters of the laser must be taken into account and be properly chosen in order to maximize efficiency.

II. Output Power from Laser Diodes

We start with the steady-state local rate equations describing the spatial evolution of the carrier and photon densities along the laser cavity (Ref 5)

$$\frac{dS^+}{dZ} = [\Gamma g(N) - \alpha_0] S^+ \quad (1a)$$

$$\frac{dS^-}{dZ} = -[\Gamma g(N) - \alpha_0] S^- \quad (1b)$$

$$0 = \frac{J}{qd} - \nu g(N) [S^+ + S^-] - \frac{N}{\tau_s} \quad (2)$$

where S^+ and S^- are the forward and backward propagating photon densities, respectively, N is the carrier density, g is the gain coefficient, which is a function of N , Γ is the mode confinement factor, α_0 is the absorption coefficient, J is the current density injected into an active region with a thickness d , q is the electron charge, τ_s is the spontaneous carrier lifetime, ν is the group velocity of the laser mode, and Z is the distance along the laser, with $Z=0$ at the mirror with the unit reflectivity

The geometry of the laser is shown in Fig 1. Its active region has a length L , width W and thickness d . It is assumed that one facet of the laser is coated so that its reflectivity is unity. The reflectivity of the other mirror is denoted by R . From symmetry considerations it is clear that the power efficiency of this laser is twice that of a laser of length $2L$ and with two mirrors of reflectivity R .

The expression for the gain g is taken as (Ref 9)

$$g = \frac{A}{\nu} (N - N_{om}) \quad (3)$$

where A is the intrinsic gain constant of the laser material and N_{om} is the minimum carrier density needed for zero gain

We postulate a solution to Eq (1) in the form

$$S^\pm(Z) = \frac{B}{A\tau_s} e^{\pm[u(Z) - \alpha Z]} \quad (4)$$

where α , the total effective loss coefficient, is given by

$$\alpha = \alpha_0 + \frac{\Gamma A N_{om}}{\nu} \quad (5)$$

B is a constant, to be determined as described below, and $u(Z)$ is a function which contains the "slow" Z -dependence of

$S^\pm(Z)$. From Eqs (1a), (3) and (4) it is easily found that

$$\frac{du}{dZ} = \frac{\Gamma A}{\nu} N(Z) \quad (6)$$

Substitution of Eqs (3) and (4) into Eq (2) yields the following expression for the carrier density

$$\frac{N(Z)}{N_{om}} = 1 + \frac{\frac{J}{J_{om}} - 1}{1 + 2B \cosh [u(Z) - \alpha Z]} \quad (7a)$$

where

$$J_{om} = \frac{qd N_{om}}{\tau_s} \quad (7b)$$

is the current density required to establish the carrier density needed for intrinsic transparency N_{om}

Substitution of Eq (7) in Eq (6), and the application of the boundary conditions $S^+(0) = S^-(0)$ and $S^-(L) = R S^+(L)$ in Eq (4) yields the following equation for $u(Z)$

$$\frac{du}{dZ} = G \left(1 + \frac{\frac{J}{J_{om}} - 1}{1 + 2B \cosh (u(Z) - \alpha Z)} \right) \quad (8a)$$

$$u(0) = 0 \quad (8b)$$

$$u(L) = \alpha L + \ln \frac{1}{\sqrt{R}} \quad (8c)$$

with

$$G = \frac{\Gamma A N_{om}}{\nu} \quad (8d)$$

Equation (8) is solved numerically. Then, given the value of $u(Z)$ the total photon density in the laser cavity can be found from

$$S_t(Z) = S^+(Z) + S^-(Z) = \frac{2B}{A\tau_s} \cosh [u(Z) - \alpha Z] \quad (9)$$

The maximum photon density is at the mirror with the reflectivity R . From Eqs (4), (8b) and (8c) we find that it is given by

$$S_{max} = S^+(L) + S^-(L) = \frac{B}{A\tau_s} \frac{1+R}{\sqrt{R}}. \quad (10)$$

Similarly, the net photon density that is emitted from the laser is given by

$$S_{out} = S^+(L) - S^-(L) = \frac{B}{A\tau_s} \frac{1-R}{\sqrt{R}} \quad (11)$$

The optical power emitted from the laser is then given by

$$P_{out} = \nu \cdot (q Eg) \cdot S_{out} \cdot W \cdot \frac{d}{\Gamma} \quad (12)$$

where it has been assumed that the energy of an emitted photon approximately corresponds to the bandgap Eg of the material.

The qualitative behavior of the photon densities, $S^+(Z)$ and $S^-(Z)$, as well as the electron density $N(Z)$ are depicted in Fig. 2

III. Power Efficiency of the Laser Diode

The first step is to find the constant B which scales the amplitudes of the photon waves (Eq 4). It is determined from either Eq (10) (for a maximum photon density constraint) or from Eq (12) (for a maximum output intensity constraint). The next step involves the numerical solution of Eq (8a) with the boundary conditions of Eqs (8b) and (8c). The solution of Eq (8a) yields the needed value of J .

The electrical power invested in the laser is given by

$$P_{elect} = J \cdot W \cdot L (Eg + \rho J) \quad (13)$$

The first term in Eq (13) is the electrical power spent at the junction itself (it is assumed that under the injection conditions occurring in laser diodes, the voltage across the junction roughly equals the bandgap), and the second term corresponds to power losses in resistive layers with resistivity of $\rho(\Omega - \text{cm}^2)$ which are in series with the junction current path.

The power efficiency η is simply found by

$$\eta = \frac{P_{out}}{P_{elect}} \quad (14)$$

where P_{out} and P_{elect} are given by Eqs (12) and (13), respectively. It should also be noted that the efficiency

obtained in Eq (14) is actually normalized to the internal efficiency η_i of the laser (Ref 10). At room temperatures η_i approximately equals 0.8.

IV. Numerical Examples and Discussion

Numerical calculations for several cases were carried out. The values of the parameters chosen were $Eg = 1.4 \text{ V}$, $\nu = 8.7 \cdot 10^7 \text{ m-sec}^{-1}$ (corresponding to an index of refraction of 3.6), $A = 1.6 \cdot 10^{-6} \text{ cm}^3\text{-sec}^{-1}$, $N_{om} = 7.5 \cdot 10^{17} \text{ cm}^{-3}$, $\tau_s = 3 \cdot 10^{-9} \text{ sec}$, $\Gamma = 0.2$, $d = 0.1 \text{ } \mu\text{m}$ and $W = 5 \text{ } \mu\text{m}$. The efficiency was calculated as a function of the laser length for several values of mirror reflectivity ($R = 0.8, 0.5, 0.32, 0.1, 10^{-2}, 10^{-3}, 10^{-4}$), absorption coefficient ($\alpha_0 = 0, 10, 20 \text{ cm}^{-1}$), laser material resistivity ($\rho = 0, 10^{-4}, 8 \cdot 10^{-4} \text{ } \Omega\text{-cm}^2$) and maximum output power intensity ($I_{max} = 10^6, 10^7 \text{ W-} \text{cm}^{-2}$).

A typical result is shown in Fig 3. In this example $\alpha = 10 \text{ cm}^{-1}$, $\rho = 10^{-4} \text{ } \Omega\text{-cm}^2$ and $I_{max} = 10^6 \text{ W-} \text{cm}^{-2}$. With these parameters, the maximum available efficiency is about 46% and it is achieved for a 125- μm -long laser with a mirror reflectivity of 0.32. The laser operates at about four times the threshold current. The total power emitted by the laser depends on the cross-sectional dimensions of its active region.

In Fig 4(a) and 4(b) we see the impact of constraints on the current density and current, respectively. We see that unless operation at $J \geq 5 \text{ kA/cm}^2$ (Fig 4(a)) or $I \geq 25 \text{ mA}$ is allowed, the maximum available efficiency cannot be obtained.

The maximum available efficiency depends heavily on the various laser parameters, as is shown in Fig 5. It is clear that the efficiency improves as α_0 and ρ decrease. It is also important to note that as we are operating the laser at higher output power levels its efficiency decreases since the optical power is proportional to the current density J , but the electrical power wasted at the resistive layers in the laser material is proportional to J^2 .

For lasers operating at high output power levels the optimum efficiency is obtained at low values of mirror reflectivity ($10^{-1} \leq R \leq 10^{-4}$). As the operating power level is decreased, however, the optimum efficiency is obtained at higher values of mirror reflectivity. This trend becomes more pronounced as α_0 and ρ decrease. Also as the power levels are decreased, the laser operates at current densities which are closer and closer to the threshold value. It can also be shown that as the mirror reflectivity becomes smaller, it takes longer lasers to obtain the optimum efficiency (Fig 6).

It is interesting to note what happens in the idealized case of $\rho = 0 \Omega\text{-cm}^2$. When $\rho = 0 \Omega\text{-cm}^2$ the efficiency is always maximized as the laser length approaches zero. This shows the importance of including ohmic losses in the analysis in spite of the fact that they are not related to the intrinsic operation of the laser.

As a final remark we will touch on the subject of the difference between the constraints on maximum output intensity or maximum photon (energy) density. For low values of mirror reflectivity there is no significant difference between the two cases because most of the energy is contained in the outward going wave. However, as the mirror reflectivity is increased, more energy is contained in the reflected wave. If the constraint is on the overall photon density it implies that we can extract less power from the laser, a fact that reduces the efficiency of the laser.

V. Conclusions

Analysis of a model of semiconductor injection lasers has been carried out in terms of the basic parameters relevant to efficient device operations (i.e., device length, mirror reflectivity, absorption coefficient and material conductivity). Not surprisingly it has been found that the efficiency is improved as the optical absorption coefficient and the resistivity of the laser material are reduced. The next step is to specify either the maximum output intensity or the maximum photon density of the laser, these parameters are constrained by reliability considerations. When all these are determined, there will exist a unique value of laser length and laser reflectivity which will optimize the laser efficiency for each application. It is important to note that all the device parameters must be adequately designed in order to optimize its overall power efficiency.

References

1. Katz, J., "High Power Semiconductor Lasers for Deep Space Communications," *TDA Progress Report 42-63*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 40-50, June 15, 1981.
2. Katz, J., and Lesh, J. R., "Optical Communication for Deep Space Missions Using Integrated Semiconductor Injection Lasers," Technical Report (in preparation), Jet Propulsion Laboratory, Pasadena, Calif.
3. Sommers, H. S., Jr., "Theoretical Maximum of the Power Efficiency of a Pulsed Injection Laser with a Fabry-Perot Cavity," *Solid-State Electron.*, **11**, pp. 909-916, 1968.
4. Whiteaway, J. E. A., and Thompson, G. H. B., "Optimization of Power Efficiency of (GaAl)As Injection Lasers Operating at High Power Levels," *IEEE J. Solid-State Electron. Dev.*, **1**, pp. 81-88, 1977.
5. Hasuo, S., and Ohmi, T., "Spatial Distribution of the Light Intensity in the Injection Lasers," *Jpn. J. Appl. Phys.*, **13**, pp. 1429-1434, 1974.
6. Yonezu, H., et al., "High Optical Power Density Emission from a 'Window-Stripe' AlGaAs Double-Heterostructure Laser," *Appl. Phys. Lett.*, **34**, pp. 637-639, 1979.
7. Henry, C. H., et al., "Catastrophic Damage of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ Double-Heterostructure Laser Material," *J. Appl. Phys.*, **50**, pp. 3721-3732, 1979.
8. Ettenberg, M., and Kressel, H., "The Reliability of (AlGa)As CW Laser Diodes," *IEEE J. Quant. Electron.*, **QE-16**, pp. 186-196, 1980.
9. Kressel, H., and Butler, J. K., *Semiconductor Lasers and Heterostructure LEDs*, Academic Press, New York (1977).
10. Yariv, A., *Quantum Electronics*, Ch 10, Wiley 1975.

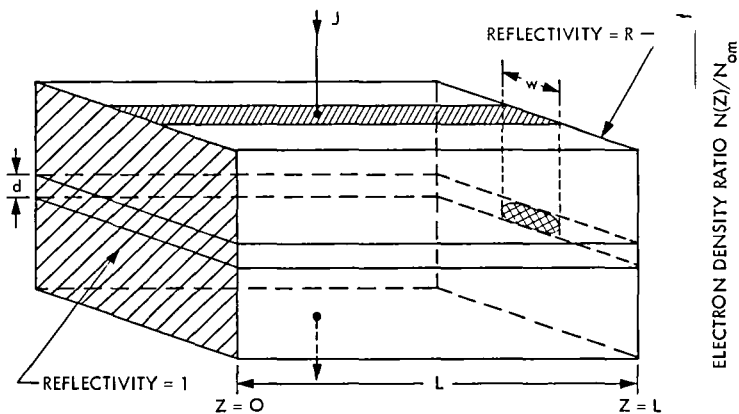


Fig 1 Schematic drawing of a semiconductor injection laser

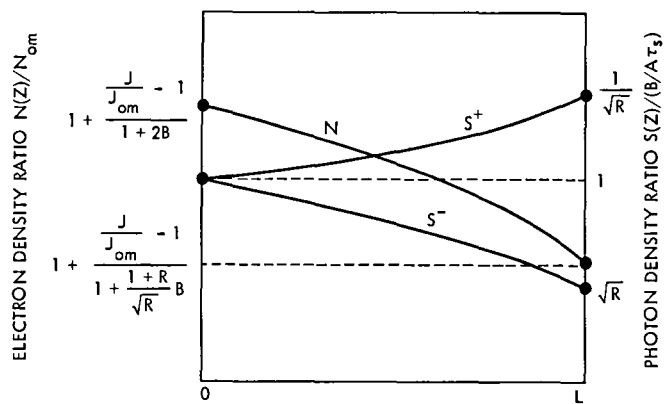


Fig 2. Spatial distribution of photon and electron densities along the laser

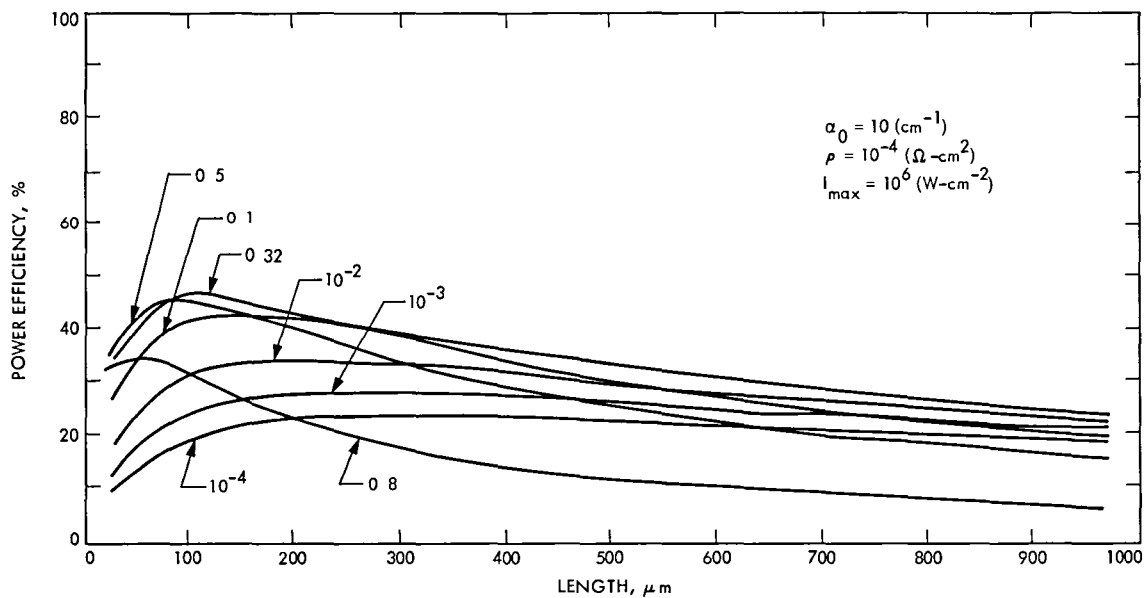


Fig 3. An example of the dependence of the laser efficiency on the laser length, with the mirror reflectivity as a parameter

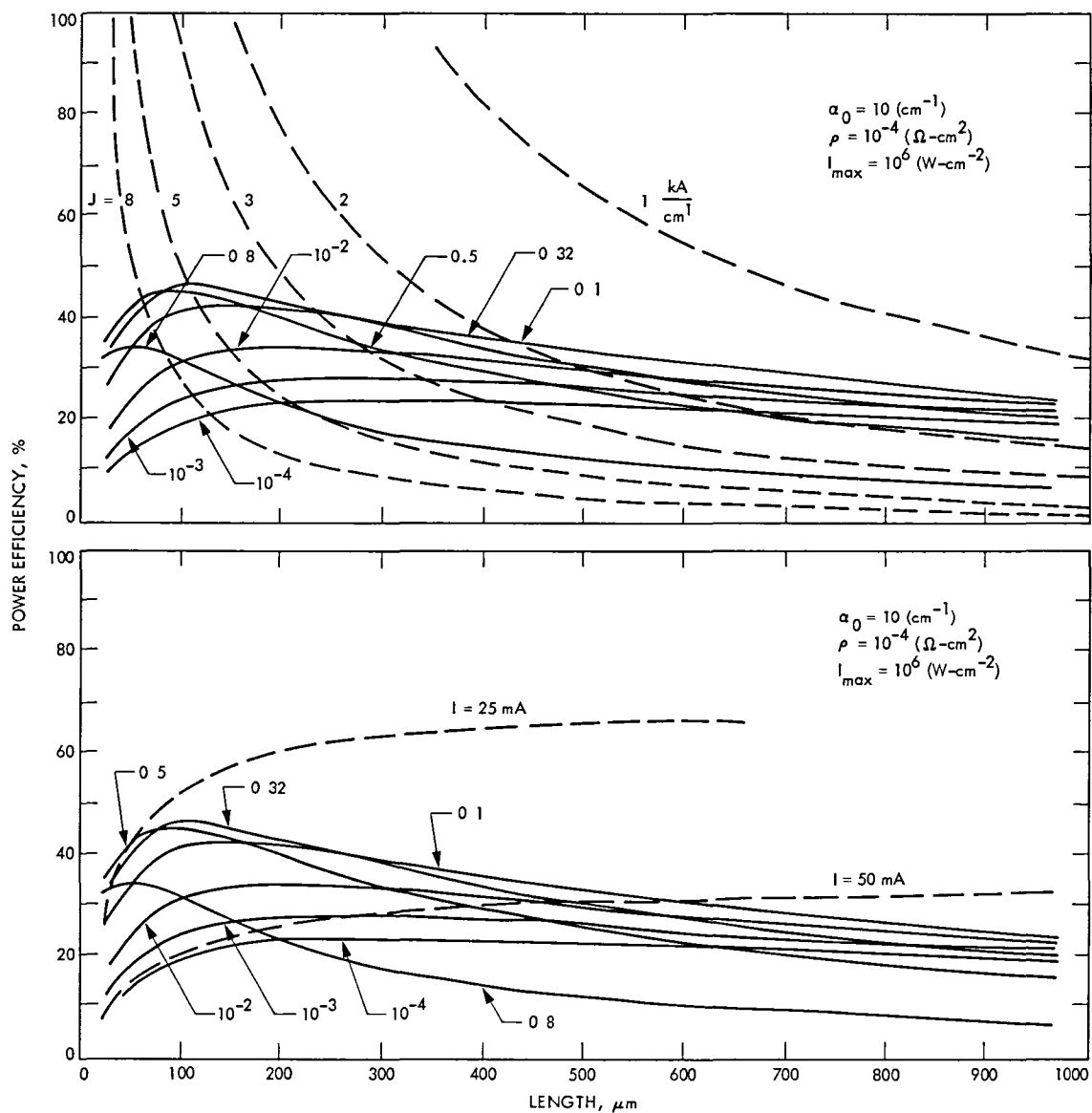


Fig. 4. Same as Fig. 3, but with dashed lines showing constraints of (a) current density and (b) current

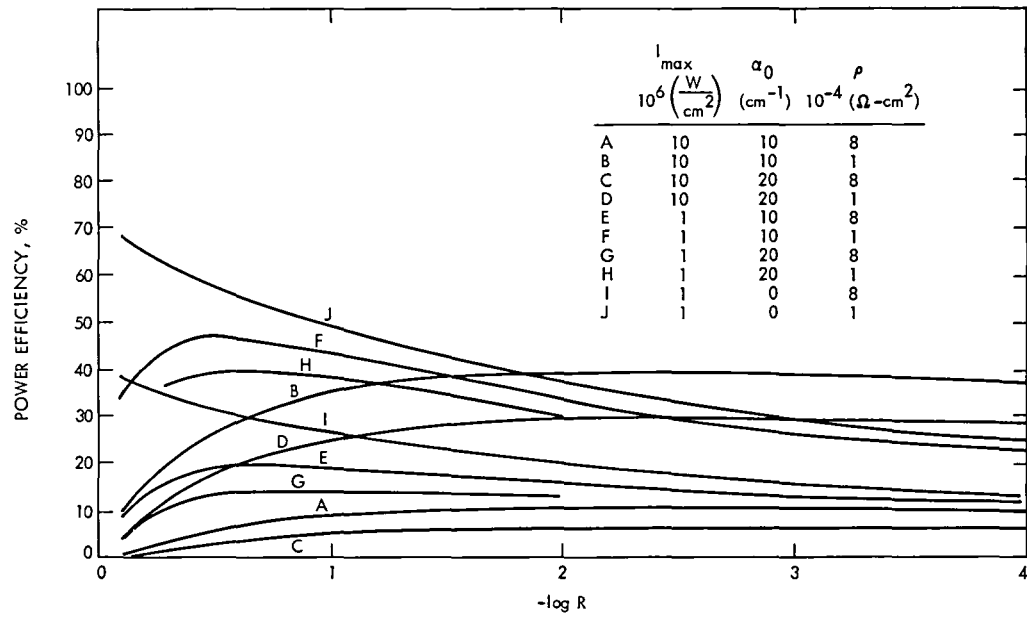


Fig. 5. Dependence of the maximum available efficiency as a function of the mirror reflectivity, with α_0 , ρ and I_{\max} as parameters

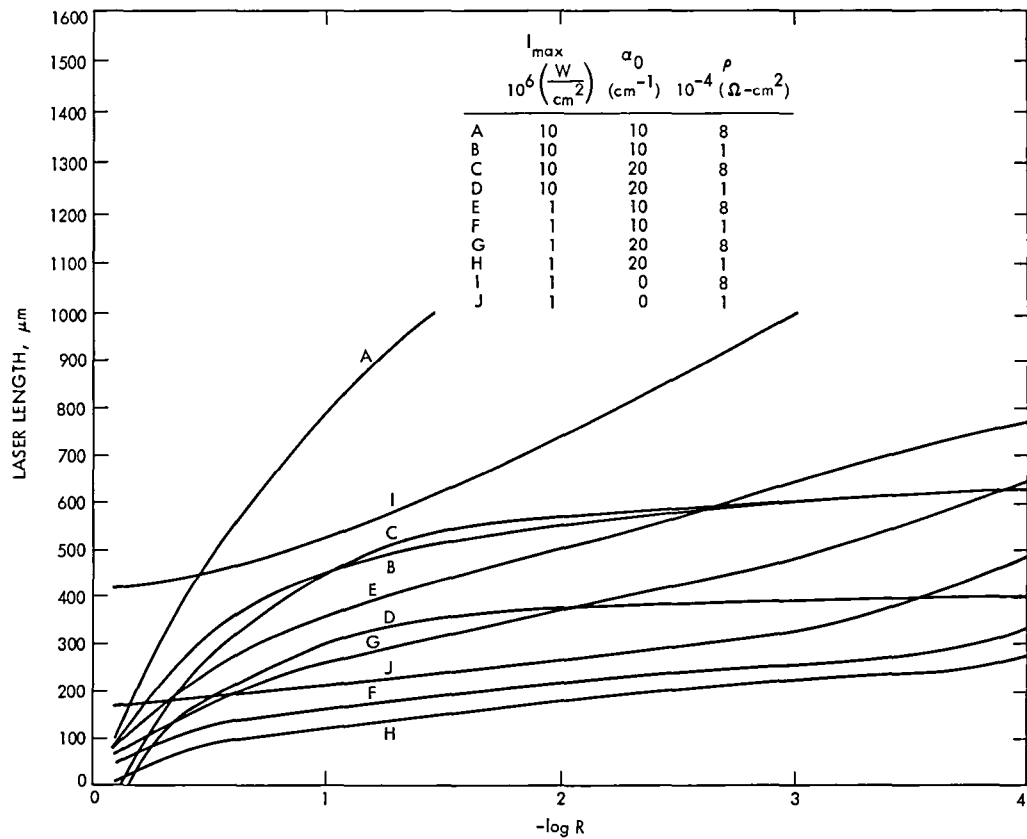


Fig. 6. Laser lengths at which the maximum efficiency is obtained (parameters are the same as in Fig. 5)

Phase-Locking of Semiconductor Injection Lasers

J. Katz

Communications Systems Research Section

Phase locking of several semiconductor injection lasers via mutual coupling is a possible method for coherent power combination. In this report the equations describing an array of semiconductor injection lasers are formulated, their solution is outlined and the conditions needed for locking (synchronization) are derived in terms of both phenomenological and actual device parameters. It is found that for real devices these conditions are at present quite stringent but not impossible to fulfill

I. Introduction

Semiconductor injection lasers are an attractive candidate for emitters in optical communication systems for deep space missions (Ref. 1). However, their advantages of small size, high reliability, good power efficiency and the possibility to directly modulate them at high rates are offset by the fact that the power levels emitted by a single device into a stable radiation pattern are too low for this application. One possible method for overcoming this problem is by the coherent combination of several semiconductor lasers, in a manner similar to a phased-array of antennas (Ref. 2).

There are several advantages of phase locking lasers. First, when the power of the lasers is combined incoherently, each laser emits light in its own individual spectrum. Thus it is necessary to have an optical filter with a wider bandwidth at the receiver, with a resulting increase in detected background noise radiation. Secondly, the locking of the laser components causes a reduction in the far-field beam divergence angle. This makes the task of subsequent beam narrowing simpler (e.g., by requiring an optical telescope with a smaller magnification).

The purpose of this report is to review the subject of mutual phase-locking (or synchronization) of semiconductor

injection lasers. In Section II the equations of motion of a single-mode individual laser are developed, using the density matrix formulation. Section III formulates the general equations for describing an array of lasers. Then some simplifying assumptions are made and the resulting working equations are developed. Section IV outlines the solution and discusses the conditions necessary for locking in terms of phenomenological coupling coefficients. Finally, Section V analyzes the coupling coefficients in different array configurations in terms of the device parameters and geometry. These results can serve as basic guidelines for implementing various array designs.

II. Equations of Motion of a Semiconductor Laser

The equations derived in this section describe the temporal evolution of the variables pertaining to the operation of an individual, single-longitudinal-mode semiconductor injection laser. All the spatial dependence is assumed to be either uniform or averaged out. The spatial mode profile will be considered later during the calculations of the various coupling coefficients, as described in Section V. The main variables describing the laser are its electric field amplitude E , the polarization of its active medium P , and its inversion density N .

The first equation to describe the laser is the wave equation of the lasing mode, which, in cgs units can be expressed as (Ref. 5)

$$\frac{d^2 E}{dt^2} + \frac{\omega}{Q_0} \frac{dE}{dt} + \omega^2 E = -4\pi \frac{d^2 P}{dt^2} \quad (1)$$

where ω is the natural frequency of the laser resonator and Q_0 is the figure of merit of the laser, given by

$$Q = \omega_0 T_{ph} \quad (2)$$

Here ω_0 is the frequency of the atomic transition, and T_{ph} is the average photon lifetime in the laser cavity.

To obtain the second equation of motion, we use the density matrix formalism of a two-level system. The semiconductor injection laser can be well approximated by a two-level system because the thermalization time within its bands is much shorter than the average interband transition times.

We start with the following equation (Ref. 3)

$$\frac{d\rho_{21}}{dt} = -i\omega_0 \rho_{21} + i \frac{\mu}{\hbar} (\rho_{11} - \rho_{22}) E - \frac{\rho_{21}}{T_2} \quad (3)$$

where ρ_{ij} is the ij th element in the density matrix, μ is the dipole moment of the transition, T_2 is the inelastic relaxation time and \hbar is Planck's constant divided by 2π . A basic introduction to the subject of density matrix formalism can be found in Ref. 3, ch. 3.

Next we write the complex conjugate equation for Eq. (3), noting that $\rho_{11} - \rho_{22}$, E , and, without loss of generality, also μ , are real variables. Adding and subtracting the equations for ρ_{21} and ρ_{21}^* , a new set of first-order differential equations for $(\rho_{21} + \rho_{21}^*)$ and $(\rho_{21} - \rho_{21}^*)$ is obtained. We can further eliminate the equation for $(\rho_{21} - \rho_{21}^*)$, thus obtaining a second-order differential equation for $(\rho_{21} + \rho_{21}^*)$. Since $\rho_{21} = \rho_{12}^*$, and, by definition

$$P = \mathcal{N} \mu (\rho_{21} + \rho_{12}) \quad (4)$$

where \mathcal{N} is the total population density, the resulting equation for $(\rho_{21} + \rho_{21}^*)$ can be rewritten as an equation for P . This equation is of the form

$$\frac{d^2 P}{dt^2} + \frac{2}{T_2} \frac{dP}{dt} + \left(\omega_0^2 + \frac{1}{T_2^2} \right) P = - \frac{2\omega_0 \mu^2}{\hbar} \mathcal{N} (\rho_{22} - \rho_{11}) E \quad (5)$$

Since we are interested in the case where $\omega_0 \gg 1/T_2$ (approximately 10^{15} vs 10^9), and the inversion density is defined by

$$N = \mathcal{N} (\rho_{22} - \rho_{11}) \quad (6)$$

then the second equation can be written in its final form

$$\frac{d^2 P}{dt^2} + \frac{2}{T_2} \frac{dP}{dt} + \omega_0^2 P = - \frac{2\omega_0 \mu}{\hbar} N E \quad (7)$$

The third and last equation of motion is derived starting with the following equation (Ref. 3)

$$\frac{d(\rho_{11} - \rho_{22})}{dt} = \frac{2i\mu E}{\hbar} (\rho_{21} - \rho_{21}^*) - \frac{(\rho_{11} - \rho_{22}) - (\rho_{11} - \rho_{22})_0}{T_1} \quad (8)$$

where T_1 is the elastic time constant of the transition and the subscript 0 denotes the steady-state solution. Using the equation for $(\rho_{21} + \rho_{21}^*)$ which was obtained during the derivation of Eq. (7), namely

$$\frac{d(\rho_{21} + \rho_{21}^*)}{dt} + \frac{\rho_{21} + \rho_{21}^*}{T_2} = -i\omega_0 (\rho_{21} - \rho_{21}^*) \quad (9)$$

in Eq. (8), multiplying the resulting equation by \mathcal{N} , and using the definitions in Eqs. (4) and (6), we obtain

$$\frac{dN}{dt} + \frac{N - N_0}{T_1} = \frac{2E}{\hbar \omega_0} \left(\frac{dP}{dt} + \frac{P}{T_2} \right) \quad (10)$$

where the steady-state influx of incoming carriers, N_0 , is defined as

$$N_0 = \mathcal{N} (\rho_{22} - \rho_{11})_0 \quad (11)$$

Since the polarization P oscillates at a frequency near ω_0 , we can assume that $dP/dt \approx \omega_0 P \gg P/T_2$, and thus we can finally write

$$\frac{dN}{dt} + \frac{N - N_0}{T_1} = \frac{2E}{\hbar\omega_0} \frac{dP}{dt} \quad (12)$$

Equations (1), (7) and (12) are the equations of motion describing the time evolution of the laser field, polarization and inversion density, respectively.

Before we continue, it is useful to convert all the variables into dimensionless parameters. The transformations are listed in Table 1, and the resulting set of dimensionless equations, equivalent to Eqs (1), (7) and (12), is

$$\frac{d^2x}{dt'^2} + b \frac{dx}{dt'} + \Omega x = -\gamma \frac{d^2v}{dt'^2} \quad (13)$$

$$\frac{d^2v}{dt'^2} + \delta_2 \frac{dv}{dt'} + v = \frac{\delta_2}{2} xw \quad (13b)$$

$$\frac{dw}{dt'} + \frac{\delta_1}{2} w = -\delta_1 - 2\delta_2 x \frac{dv}{dt'} \quad (13c)$$

Note that the time derivatives are now taken with respect to the dimensionless time $t' = \omega_0 t$, and that δ_1 , δ_2 , b , γ and Ω are much smaller than unity.

III. Equations Describing an Array of Lasers

In this section we consider the case of M semiconductor injection lasers which we coupled among themselves. Since the magnitude of the coupling is small under virtually all practical situations, we need to consider only interactions between each laser in the array and its closest neighbors. Furthermore, if we assume that we have an array which is arranged in a one-dimensional configuration, then each laser has two closest neighbors, except those at the extreme locations which have only one. Extension of the following analysis to the more general case is straightforward, although the calculations become more cumbersome.

In light of the above, Eqs. (13) are modified as follows in order to include the coupling terms (Ref. 4)

$$\begin{aligned} \ddot{x}_i(t') + b_i \dot{x}_i(t') + \xi_{i,i-1} \dot{x}_{i-1}(t' - \tau'_{i,i-1}) \\ + \xi_{i,i+1} \dot{x}_{i+1}(t' - \tau'_{i,i+1}) + \Omega_i x_i(t') = -\gamma_i \ddot{v}_i(t') \end{aligned} \quad (14a)$$

$$\ddot{v}_i(t') + \delta_2 \dot{v}_i(t') + v_i(t') = \frac{\delta_2}{2} w_i(t')$$

$$\begin{aligned} [x_i(t') + \eta_{i,i-1} x_{i-1}(t' - \tau'_{i,i-1}) \\ + \eta_{i,i+1} x_{i+1}(t' - \tau'_{i,i+1})] \end{aligned} \quad (14b)$$

$$\dot{w}_i(t') + \frac{\delta_1}{2} w_i(t') = -\delta_1 - 2\delta_2 \dot{v}_i(t')$$

$$\begin{aligned} [x_i(t') + \eta_{i,i-1} x_{i-1}(t' - \tau'_{i,i-1}) \\ + \eta_{i,i+1} x_{i+1}(t' - \tau'_{i,i+1})] \end{aligned} \quad (14c)$$

where the subscripts $i, i-1, i+1$ index the lasers in the array and the dots denote differentiation with respect to the dimensionless time t' , ξ_{ij} and η_{ij} are parameters associated with the coupling between the i th and the j th laser via their electric fields, and τ'_{ij} is the effective dimensionless delay time in the field coupling between the i th and the j th lasers

It is important to note that lasers that are located at the edge of the array (e.g., $i=1$) interact only with one other laser (since, for example x_0 does not exist). We can also see how Eqs (14) can be extended to the more general configurations, simply by adding the effects of other lasers as needed to each equation, with the use of the proper coupling coefficients. Several conditions can be assumed in order to simplify the problem.

First, if the delay time is short enough, it can be neglected. This condition is formulated as follows (Ref. 4)

$$\omega_0 \tau_{ij} = \tau'_{ij} \ll \frac{1}{\sqrt{\xi_{ij}^2 - \frac{(\omega_i - \omega_j)^2}{\omega_0^2}}} \quad (15)$$

or, equivalently

$$\tau_{ij} \ll \frac{1}{\sqrt{(\omega_0 \xi_{ij})^2 - (\omega_i - \omega_j)^2}} \quad (16)$$

The worst case in satisfying the condition in Eq (16) is when $\omega_i = \omega_j$. Assuming this and also that

$$\tau_{ij} \cong \frac{d_{ij}}{c/n} \quad (17)$$

where d_{ij} is the distance between the i th and j th lasers and n is the index of refraction of the medium, and noting that $\omega_0 = 2\pi c/\lambda$ (λ is the vacuum wavelength), the condition for negligible delay time is

$$\xi_{ij} \ll \frac{\lambda}{2\pi d_{ij} n} \quad (18)$$

which is satisfied under almost all practical situations.

The second simplification involves neglecting the coupling between the laser fields via the active medium. This condition is formulated as

$$\delta_2 \eta_{ij} \ll \xi_{ij} \quad (19)$$

In the following it will be assumed that the conditions expressed in Eqs (18) and (19) are fulfilled. The consequences of violating them will be mentioned in the next section. It will also be assumed that the elements of the array are identical and equispaced, so that $\xi_{i,i\pm1} = \xi$, $b_i = b$, and $\gamma_i = \gamma$.

Under all the above assumptions, the working formulas for describing an array of coupled lasers are

$$\ddot{x}_i(t') + b\dot{x}_i(t') + \xi [\dot{x}_{i-1}(t') + \dot{x}_{i+1}(t')] + \Omega_i x_i(t') = -\gamma \dot{v}_i(t') \quad (20a)$$

$$\ddot{v}_i(t') + \delta_2 \dot{v}_i(t') + v_i(t') = \frac{\delta_2}{2} w_i(t') x_i(t') \quad (20b)$$

$$\dot{w}_i(t') + \frac{\delta_1}{2} w_i(t') = -\frac{\delta_1}{2} - 2\delta_2 \dot{v}_i(t') x_i(t') \quad (20c)$$

IV. Solution of the Equations

For solving the set of Eqs. (20), the "fast" temporal behavior is first factored out, i.e., we assume that

$$x_i = X_i \cos(t' + \phi_i) \quad (21a)$$

$$v_i = V_i \cos(t' + \psi_i) \quad (21b)$$

$$w_i = W_i = \text{const.} \quad (21c)$$

where the last condition implies that the spectrum of the laser driving signal lies well below its resonant frequency, which is typically a few gigahertz. Substituting Eqs (21) into Eqs (20), results in, after a considerable amount of algebraic manipulation and keeping only first-order terms under consideration,

the following set of first-order nonlinear differential equations for the "slow" amplitudes of the variables

$$\begin{aligned} \dot{X}_i = & -\frac{b}{2} X_i - \frac{\xi}{2} [X_{i+1} \cos(\phi_{i+1} - \phi_i) \\ & + X_{i-1} \cos(\phi_{i-1} - \phi_i)] + \frac{\gamma V_i}{2} \sin(\psi_i - \phi_i) \end{aligned} \quad (22a)$$

$$\begin{aligned} \dot{\phi}_i = & \frac{1}{2} (\Omega_i - 1) - \frac{\xi}{2X_i} [X_{i+1} \sin(\phi_{i+1} - \phi_i) \\ & + X_{i-1} \sin(\phi_{i-1} - \phi_i)] - \frac{\gamma}{2} \frac{V_i}{X_i} \cos(\psi_i - \phi_i) \end{aligned} \quad (22b)$$

$$\dot{V}_i = -\frac{\delta_2 V_i}{2} - \frac{\delta_2}{4} X_i W_i \sin(\psi_i - \phi_i) \quad (22c)$$

$$\dot{\psi}_i = -\frac{\delta_2}{4} \frac{X_i W_i}{V_i} \cos(\psi_i - \phi_i) \quad (22d)$$

$$\dot{W}_i = -\frac{\delta_1}{2} W_i - \delta_1 + \delta_2 X_i V_i \sin(\psi_i - \phi_i) \quad (22e)$$

The first subject to be investigated is the steady-state solution of the above equations. For this we let $\dot{X}_i = \dot{V}_i = \dot{W}_i = 0$ (i.e., constant amplitudes) and $\dot{\phi}_i = \dot{\psi}_i \equiv \delta\omega$ (i.e., phase locking) in Eqs (22). The resulting set of nonlinear algebraic equations describing the steady-state array variables is given by

$$\begin{aligned} X_i \left[b - \frac{\gamma \sin^2(\psi_i - \phi_i)}{1 + \frac{\delta_2}{\delta_1} X_i^2 \sin^2(\psi_i - \phi_i)} \right] = \\ -\xi [X_{i+1} \cos(\phi_{i+1} - \phi_i) + X_{i-1} \cos(\phi_{i-1} - \phi_i)] \end{aligned} \quad (23a)$$

$$\begin{aligned} 2\delta\omega + (1 - \Omega_i) + \gamma \frac{\sin(\psi_i - \phi_i) \cos(\psi_i - \phi_i)}{1 + \frac{\delta_2}{\delta_1} X_i^2 \sin^2(\psi_i - \phi_i)} = \\ -\frac{\xi}{X_i} [X_{i+1} \sin(\phi_{i+1} - \phi_i) \\ + X_{i-1} \sin(\phi_{i-1} - \phi_i)] \end{aligned} \quad (23b)$$

$$\delta\omega = \frac{\delta_2}{2} \cos(\psi_i - \phi_i) \quad (23c)$$

Since one of the ϕ_i 's (or ψ_i 's) can be set arbitrarily to zero, and in all the lasers $(\psi_i - \phi_i)$ is the same, Eqs (23) is a set of $2M$ equations with $2M$ unknowns, where M is the number of lasers in the array. Once Eqs. (23) are solved (which can be done only numerically), the other variables can be expressed in terms of the X_i 's, ϕ_i 's and ψ_i 's, as follows

$$W_i = - \frac{2}{1 + \frac{\delta_2}{\delta_1} X_i^2 \sin^2(\psi_i - \phi_i)} \quad (24)$$

and

$$V_i = \frac{X_i \sin(\psi_i - \phi_i)}{1 + \frac{\delta_2}{\delta_1} [X_i \sin(\psi_i - \phi_i)]^2} \quad (25)$$

The uncoupled (i.e., $\xi = 0$) solution of Eqs (23) – (25) is obtained straightforwardly. In this case the phase shifts between the ϕ_i 's are undefined, the phase shift in each laser between its field and its polarization is

$$\tan(\psi_i^0 - \phi_i^0) = \left[\frac{b + \delta_2}{\Omega_i - 1} \right] \quad (26)$$

approaching $\pi/2$ as the natural frequency of the resonator is closer to the transition frequency, and the field amplitude is given by

$$|X_i^0| = \sqrt{\frac{\delta_1}{\delta_2}} \sqrt{\frac{\gamma}{b} - \left(\frac{\Omega_i - 1}{b + \delta_2} \right)^2 - 1} \quad (27)$$

where the pumping term is contained in γ (see Table 1). In Appendix A the relation between the formalism used in this report and the solutions obtained from the laser rate equations is discussed

Since the phenomenon of locking is a nonlinear one, we cannot solve for the coupled case by using perturbation methods on the uncoupled case, especially since the phase relationships between the lasers have a sudden transition as locking occurs. However, several other numerical methods, such as successive iterations, can be used to yield the desired results

If we sum up Eq (23b) for $i = 1, 2, \dots, M$, we find that if ω_0 is the “center of mass” of all the uncoupled lasers’ frequencies, i.e.,

$$\sum_{i=1}^m X_i^2 (1 - \Omega_i) = 0 \quad (28)$$

then we have

$$\psi_i - \phi_i = \frac{\pi}{2} \quad (29)$$

and the synchronized array oscillates at the center transition frequency ω_0 .

As shown in Refs. 4–6, all but one of the conditions for stability are identical to the conditions of existence of a solution. The stability investigation is done on the equations of the “slow” variables in Eqs (22), with the derivative of the “fast” variables set to zero (since they are “fast” they can always follow adiabatically the “slow” variables). For semiconductor injection lasers with common parameters, the following conditions hold

$$T_2, T_{ph} \gg T_1 \quad (30)$$

and thus W_i and X_i are the “slow” variables and V_i are the “fast” variables. An exact stability analysis of this nonlinear problem is prohibitively complicated, so a linearized sensitivity test about the steady-state solution is carried out, as outlined in Refs 4 and 5.

To conclude, we can say that if a solution exists, and if the figure of merit of the laser resonator increases when becoming part of the array, the solution is stable. In the following, these stability conditions will be stated more explicitly.

Successive application of Eq (23b), using Eqs (28) and (29), yields the following expression for the sine of the phase shift between two neighboring lasers

$$\sin(\phi_{k+1} - \phi_k) = - \frac{1}{\xi X_{k+1} X_k} \sum_{i=1}^k X_i^2 (1 - \Omega_i) \quad (31)$$

Since the sine value of a real angle is always bounded by unity, the right-hand side of Eq (31) must be bounded by unity as expressed in the following equation

$$\left| \sum_{i=1}^k X_i^2 (1 - \Omega_i) \right| < \left| \xi X_{k+1} X_k \right|$$

$$i = 1, 2, \dots (M-1) \quad (32)$$

When the coupling is not too strong, then to a first approximation $X_1 \cong X_2 \cong \dots X_M$, and the condition in Eq (32) simplifies to

$$\left| \sum_{i=1}^k (1 - \Omega_i) \right| < \left| \xi \right| \quad i = 1, 2, \dots (M-1) \quad (33)$$

The second stability condition requires that the intensity of the field (i.e., X^2) is to be a positive number. From Eq (23a) (using again Eqs. (28) and (29)) this condition can be expressed as

$$X_k \frac{\gamma - b}{\xi} > X_{k+1} \cos(\phi_{k+1} - \phi_k) + X_{k-1} \cos(\phi_k - \phi_{k-1})$$

$$k = 1, 2, \dots M \quad (34)$$

or, using Eq. (31)

$$X_k^2 (\gamma - b) > \sqrt{(\xi X_{k+1} X_k)^2 - \left[\sum_{i=1}^k X_i^2 (1 - \Omega_i) \right]^2}$$

$$+ \sqrt{(\xi X_k X_{k-1})^2 - \left[\sum_{i=1}^{k-1} X_i^2 (1 - \Omega_i) \right]^2}$$

$$k = 1, 2, \dots M \quad (35)$$

Again, this condition simplifies considerably in the region where the coupling is not strong, and thus $X_1 \cong X_2 \cong \dots X_M$. In that case the condition in Eq. (35) simplifies to

$$\gamma - b > \sqrt{\xi^2 - \left[\sum_{i=1}^k (1 - \Omega_i) \right]^2} + \sqrt{\xi^2 - \left[\sum_{i=1}^{k-1} (1 - \Omega_i) \right]^2}$$

$$k = 1, 2, \dots m. \quad (36)$$

The last stability condition comes from energetic consideration. The laser needs to "gain" something by "joining" the array, i.e., its Q needs to increase. This implies that

$$\xi < 0 \quad (37)$$

Equations (33), (36) and (37) summarize the approximate conditions for existence and stability of a solution for an array of semiconductor lasers. It should be noted, however, that as the coupling constant ξ becomes larger and approaches the value of $b (= 1/Q)$, these approximations cease to be accurate and Eqs (32) and (35) have to be used instead of Eq (33) and (36).

The reason is that when ξ is comparable to b , the change in $1/Q$ due to the external interaction is comparable to the intrinsic $1/Q$ of the resonator. In this limit of very strong coupling, each individual laser loses its identity almost completely, and the intensities of each laser vary considerably, being strongly influenced by the phase-shifts among the lasers. In this region it is more relevant to analyze the whole array as one "super waveguide" (Ref 7). In most practical cases, however, the limit of very strong coupling is not reached.

V. Coupling Coefficients in Different Configurations

In the preceding section the conditions for obtaining phase locking of an array of lasers were derived in terms of phenomenological coupling parameters (ξ). In this section we will relate these coupling parameters to the actual device parameters. Most of the section will discuss mutual coupling between lasers via field interaction due to their close proximity. The subject of diffraction coupling will also be briefly discussed.

Coupling between two semiconductor lasers which are in close proximity to each other is equivalent to coupling between two waveguides (see Fig 1). The coupling strength commonly used in the literature is defined as coupling per unit length and is denoted by K (cm^{-1}). We first want to establish the relation between K and the coupling parameter ξ used in the wave equation. By inspection of Eq. (14a) we see that ξ_{ij} is the change in the effective $1/Q$ of the i th laser in the array due to its interaction with the j th laser. The fraction of the field lost during one optical cycle is π/Q , so the fractional field coupled during the time of one optical cycle is $\pi\xi$. The fraction of the field coupled when the wave propagates a distance of 1 cm is $|K|$. The wave travels this distance in a time that is equivalent to n/λ optical cycles, where λ is the vacuum wavelength of the radiation and n is the index of refraction of the material.

Thus, we can conclude that

$$\pi\xi = \frac{\lambda}{n} K \quad (38)$$

or

$$\xi = \frac{2K}{k} \quad (39)$$

where $k = 2\pi n/\lambda$.

In order to calculate the coupling coefficient K between two semiconductor lasers in close proximity, we must solve first for the eigenmodes and the eigenvalues of a waveguide which is made of a region with an index of refraction n_2 which is surrounded by two semi-infinite regions with an index of refraction n_1 ($n_1 < n_2$). Such a waveguide is depicted in Fig. 2. By using the effective index formalism (Ref. 9) we can reduce the two-dimensional problem of a laser waveguide cross-section (see Fig. 1) to the one-dimensional waveguide depicted in Fig. 2

The transverse profile of the eigenmodes field is given by

$$E(x) = \begin{cases} \cos(hx) & |x| < \frac{s}{2} \\ \cos\left(\frac{hs}{2}\right) e^{q\frac{s}{2}} e^{-q|x|} & |x| > \frac{s}{2} \end{cases} \quad (40a)$$

$$\cos\left(\frac{hs}{2}\right) e^{q\frac{s}{2}} e^{-q|x|} \quad |x| > \frac{s}{2} \quad (40b)$$

where

$$h^2 \equiv n_2^2 k_0^2 - \beta^2 \quad (41a)$$

$$q^2 \equiv \beta^2 - n_1^2 k_0^2 \quad (41b)$$

and β is the propagation constant eigenvalue of the mode (see below). The general solution for the eigenvalue involves a transcendental equation (Ref. 3, ch. 19). However, an approximate analytical solution is possible in the following two cases

(1) Weakly guided modes, i.e.,

$$n_2^2 - n_1^2 \ll \left(\frac{\lambda}{2s}\right)^2 \quad (42)$$

In this case

$$h \cong k \sqrt{n_2^2 - n_1^2} \quad (43a)$$

$$q \cong \frac{(n_2^2 - n_1^2) k^2 s}{2} \quad (43b)$$

$$\beta \approx n_1 k \quad (43c)$$

$$q \ll h \ll \frac{\pi}{s} \quad (43d)$$

(2) Well-confined modes, i.e.,

$$n_2^2 - n_1^2 \gg \left(\frac{\lambda}{2s}\right)^2 \quad (44)$$

In this case

$$h \cong \frac{\pi}{s} \quad (45a)$$

$$q \cong k \sqrt{n_2^2 - n_1^2} \quad (45b)$$

$$\beta \cong n_2 k \quad (45c)$$

$$q \gg h \cong \frac{\pi}{s} \quad (45d)$$

Next we assume that the field propagating in the lasers in one direction is of the form

$$\mathcal{E}(x, z, t) = E(x, z) e^{i(\omega t - \beta z)} \quad (46)$$

This leads, after using the usual adiabatic approximation, to the following wave equation

$$\frac{d^2 E}{dx^2} = 2i\beta \frac{dE}{dz} + (\epsilon k^2 - \beta^2) E = 0 \quad (47)$$

We assume that the field growth along the Z direction is due to net gain g in its own active medium as well as the coupling K from the field of the other laser, and that the two interacting waves propagate codirectionally. The maximum coupling occurs when the propagation constants in the two lasers are identical (i.e., phase-matching), in which case we can write (Ref. 3, ch. 16)

$$\frac{dE_1}{dz} = gE_1 - iK E_2 \quad (48)$$

We define the complex index of refraction as

$$\tilde{\epsilon} = \epsilon + 2i \frac{\beta}{k^2} g \quad (49)$$

Using Eqs (48) and (49) or (47) results in the following equation for the transverse profile of the field

$$\frac{d^2 E_1}{dx^2} + (\tilde{\epsilon} k^2 - \beta^2) E_1 + 2 \beta K E_2 = 0 \quad (50)$$

Multiplying Eq. (50) by E_2^* and integrating (also integration by parts) yields the following expression for K

$$K = \frac{\int \left(\frac{dE_1}{dx} \right) \left(\frac{dE_2}{dx} \right)^* dx - \int (\tilde{\epsilon} k^2 - \beta^2) E_1 E_2^* dx}{2 \beta \int |E_2|^2 dx} \quad (51)$$

Since the coupling is small, we can take for E_1 and E_2 the unperturbed field profiles, as given by Eq (40) for E_1 and its displaced version for E_2 .

For the case of weakly guided modes (Eq. (42)) we obtain

$$K = - \frac{\left[s (n_2^2 - n_1^2) \right]^2 k^3}{8 \sqrt{\epsilon}} e^{-\frac{s(n_2^2 - n_1^2) k^2 d}{2}} \quad (52)$$

where d is the separation between the two lasers (see Fig 1) and $\sqrt{\epsilon} = n_2 \cong n_1$. Note that K , and hence ξ , are negative, as required (Eq. (37))

This coupling has a maximum when

$$s(n_2^2 - n_1^2) = \frac{4}{k^2 d} \quad (53)$$

and the resulting optimum coupling is

$$|K|_{\max} = \frac{\lambda_0}{\pi e^2 d^2 \sqrt{\epsilon}} \quad (54)$$

Using the relationship between K and ξ (Eq 30), we can write

$$|\xi|_{\max} = \left(\frac{2}{\sqrt{\epsilon} \pi e d} \right)^2 \quad (55a)$$

If we take GaAs as a typical example, then we have $\lambda = 0.9 \mu\text{m}$ and $\sqrt{\epsilon} \cong 3.6$, in which case we obtain

$$|\xi|_{\max}^{(\text{GaAs})} \cong 8 \cdot 10^{-4} \frac{1}{d^2} \quad (55b)$$

where d is expressed in micrometers

For well-confined modes (Eq 36), the coupling is given by

$$|K| = \frac{\lambda^2}{2 \sqrt{\epsilon} \sqrt{n_2^2 - n_1^2} s^3} e^{-k d \sqrt{n_2^2 - n_1^2}} \quad (56)$$

and is usually much smaller than the coupling in the other case (Eqs 52, 54).

As an example of the application of Eq. (55) to the design of laser arrays, let us assume that the differences in the natural oscillation frequencies of the individual lasers are due to different lengths

$$\frac{(1 - \Omega_i)}{2} = \frac{\Delta \omega_i}{\omega_0} = - \frac{\Delta L_i}{L_i} \quad (57)$$

where L_i is the length of the cavity of the i th laser. From Eqs (33), (55) and (57) we find that

$$\left| \sum_{i=1}^k \Delta L_i \right| \lesssim 5 \cdot 10^{-4} L \left(\frac{\lambda}{d} \right)^2 \quad i = 1, 2, \dots, M \quad (58)$$

when $L_i \cong L_2 \cong L_1 = L$. For $L \cong 300 \mu\text{m}$ and $\lambda/d \cong 0.1$, we obtain

$$\left| \sum_{i=1}^k \Delta L_i \right| \lesssim 15 \text{ \AA}$$

which is a rather stringent requirement. For $(\lambda/d) \sim 0.2$, the above constraint is increased to 60 \AA.

We can now check the validity of the approximation of Eq (19) by neglecting the coupling of the fields via the active medium. Assuming weakly guided modes (which provide for stronger coupling) and that a uniform carrier profile exists at the laser stripe of a width s as well as an additional region of width $s_0/2$ on both of its sides, we obtain

$$\eta = q(s + s_0) e^{-q d} \quad (59)$$

Using Eqs (33), (43b) and (52), we see that the condition expressed in Eq (19) is fulfilled if the following condition holds

$$\frac{\Delta n}{n} \gg \delta_2 \quad (60)$$

where $\Delta n = n_2 - n_1$. The condition implied in Eq (60) states that the fractional change of the index of refraction across the laser structure is much larger than the normalized spontaneous carrier lifetime, and this condition holds in most cases. If necessary, the effect of η can be included by replacing $|\xi|$ by $|\xi| + \delta_2|\eta|$ in all the stability conditions. We see that the increased coupling via the active region allows phase-matching for greater frequency deviations.

Finally, we will briefly summarize the case where we have diffraction coupling. The lasers are put in an external cavity, and part of the field of each laser is reflected into the other lasers producing a coupling between them. The magnitude of the coupling coefficient in this case has been determined in Refs. 5, 6, and 8. The general calculations in this case do not yield simple analytical expressions (Refs 5, 8). However, if we assume that the sides of the lasers facing the external mirror are coated with an anti-reflection coating and that neighboring lasers can interact with one another, i.e.,

$$\theta_{F \cdot F} > \frac{d + s}{L_{ext}} \quad (61)$$

where $\theta_{F \cdot F}$ is the far-field angle of the laser radiation pattern in the plane of the array configuration, and L_{ext} is the distance from the end face of the laser to the external mirror (Fig. 3), then the coupling constant is given by (Ref. 6)

$$\xi_{diff} = -\frac{1}{2} \left(\frac{\lambda}{2\pi} \right)^3 \frac{1}{Ls(s+d)(n_2^2 - n_1^2)} \quad (62)$$

where L is the total length of the resonator. The coupling constant is inversely proportional to the distance between the

lasers, to the length of the resonator and to the stripe width of each laser. The last functional relationship is due to the fact that the far-field angle of the laser radiation is inversely proportional to its stripe width.

Dividing Eqs. (55) and (62) we find that the ratio between the two is approximately

$$\frac{\xi_{fields\ overlap}}{\xi_{diffraction}} = \left(\frac{L}{\lambda/n} \right) \frac{s}{d} \left(1 + \frac{s}{d} \right) \Delta n \quad (63)$$

Depending on the array parameters, one of the two coupling mechanisms is dominant. For example, from Eq (63) we see that as the separation d between the individual lasers increases, there is more coupling due to diffraction than due to field overlap (note, however, that both couplings are at least inversely proportional to d).

VI. Conclusions

The problem of phase-locking of several semiconductor injection lasers via mutual coupling has been investigated. First, the equations of motion of the parameters of both individual lasers and an array of lasers were derived, considering both coupling via field overlapping and interactions via the laser active media. The solution of the equations was outlined and, more important, the conditions for obtaining phase-matching were derived. These conditions put restrictions on both the sign and the magnitude of the coupling constraint: it must be negative so that phase-matched (synchronized) operation is energetically favorable, and its magnitude must be greater than the standard deviation of the lasers' frequency deviations from the laser transition frequency.

Finally, the actual magnitudes of the coupling coefficients in several configurations were calculated, and the results indicate that it is possible, although not trivial, to achieve the coupling conditions.

References

1. Katz, J , and Lesh, J. R., "Optical Communications for Deep Space Missions Using Integrated Semiconductor Injection Lasers," Jet Propulsion Laboratory, Pasadena, Calif (in preparation)
2. Katz, J , "High Power Semiconductor Lasers for Deep Space Communications," *TDA Progress Report 42-63*, pp. 40-50, Jet Propulsion Laboratory, Pasadena, Calif , June 15, 1981.
3. Yariv, A , *Quantum Electronics*, 2nd edition, ch 8, Wiley, New York, 1975
4. Belenov, E. M , and Letokhov, V S , "Theory of Coupled Optical Masers," *Sov Phys Tech Phys* , 10, pp. 1628-1629, 1966
5. Basov, N. G , Belenov, E M., and Letokhov, V. S., "Diffraction Synchronization of Lasers," *Sov Phys Tech. Phys.*, 10, pp 845-850, 1965
6. Basov, N G , Belenov, E M , and Letokhov, V S , "Synchronization of Oscillations in a Semiconducting Laser Having Several p-n Junctions," *Sov Phys. Solid State*, 7, pp 275-276, 1965.
7. Yeh, P , Yariv, A., and Hong, C. S., "Electromagnetic Propagation in Periodic Stratified Media I General Theory," *J Opt. Soc Am.* 67, pp. 423-438, 1977
8. Letokhov, V. S , "Diffraction Losses of an Open Resonator Formed by Mirrors with Absorbing Strips," *Sov Phys. Tech. Phys* , 10, pp 629-632, 1965.
9. See, for example, W T. Tsang and R. A. Logan, "GaAs-Al_xGa_{1-x}As Strip Buried Heterostructure Laser," *IEEE J. Quant. Electron QE-15*, pp 451-469, 1979.
10. Kressel, H., and Butler, J. K., *Semiconductor Lasers and Heterojunction LEDs*, Ch. 17, Academic Press, New York, 1977.

Table 1 List of transformations into dimensionless parameters

	Old variable	New dimensionless parameter	Transformation
Time	t	t'	$t' = \omega_0 t$
Electric field	E	x	$x = \frac{\mu T_2}{\hbar} E$
Polarization	P	ν	$\nu = \frac{1}{\mu N_0} P$
Inversion density	N	w	$w = -\frac{2}{N_0} N$
Pump density	N_0	γ	$\gamma = \frac{4\pi\mu^2}{\hbar} T_2 N_0$
Elastic rate of the atomic transition (transition linewidth)	T_1^{-1}	δ_1	$\delta_1 = \frac{2}{\omega_0 T_1}$
Inelastic rate of the atomic transition (corresponds to the carrier's lifetime)	T_2^{-1}	δ_2	$\delta_2 = \frac{2}{\omega_0 T_2}$
Figure of merit of laser resonator	Q	b	$b = \frac{1}{Q} \left(\frac{\omega}{\omega_0} \right)$
Frequency of laser oscillation	ω	Ω	$\Omega = \left(\frac{\omega}{\omega_0} \right)^2$

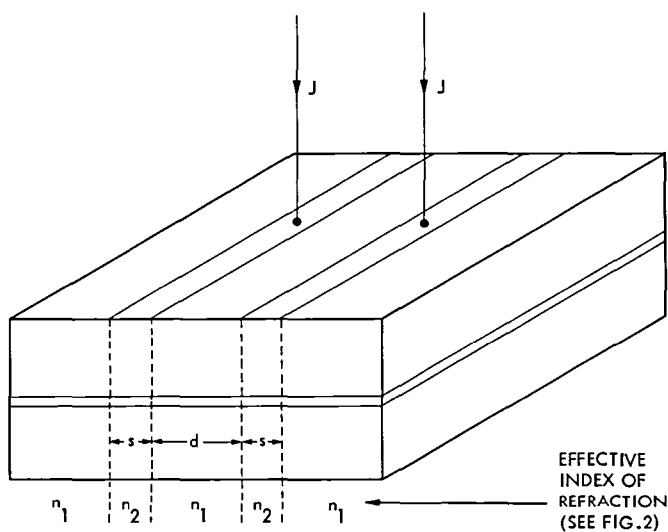


Fig 1. Schematic configuration of two semiconductor lasers in close proximity

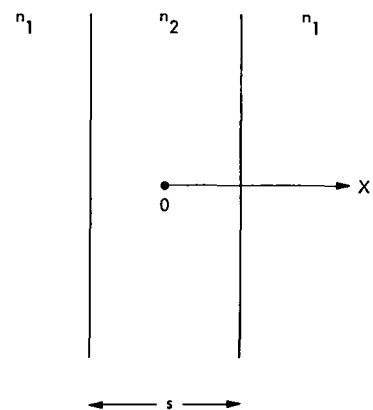


Fig 2. Schematic drawing of a one-dimensional symmetric slab waveguide

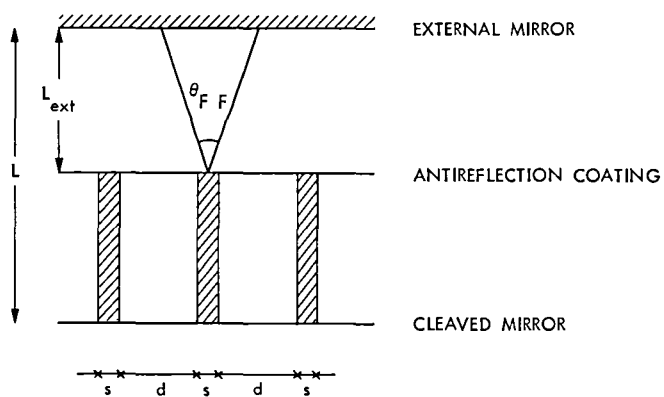


Fig 3. Schematic configuration of several semiconductor lasers in an external resonator

Appendix A

Relation Between Parameters in this Article and Parameters Appearing in the Laser Rate Equations

The linearized steady-state rate equations can be written as (Ref. 10)

$$0 = \frac{J}{qd} - ANS - \frac{N}{\tau_s} \quad (\text{A-1})$$

$$0 = ANS - \frac{S}{\tau_{ph}} \quad (\text{A-2})$$

where N and S are the carrier and photon densities, respectively, τ_s and τ_{ph} are the carrier and photon lifetimes, respectively, A is the gain coefficient, J is the current density, q is the electron charge and d is the thickness of the active region.

Solution of Eqs (A-1), (A-2) yield

$$N = \frac{\frac{J\tau_s}{qd}}{1 + A\tau_s S} \quad (\text{A-3})$$

$$A\tau_s S = \frac{JA\tau_s\tau_{ph}}{qd} - 1 \quad (\text{A-4})$$

From Eqs (24), (29) and Table 1 we obtain

$$N = \frac{N_0}{1 + \frac{T_1}{T_2} X^2} \quad (\text{A-5})$$

Thus, we can identify N_0 with

$$N_0 = \frac{J\tau_s}{qd} \quad (\text{A-6})$$

and

$$\frac{T_1}{T_2} X^2 = A\tau_s S \quad (\text{A-7})$$

Since by definition $X^2 = E^2/\hbar\omega_0$ and $T_2 = \tau_s$, we also obtain the following relations

$$A = 4\pi \frac{\mu^2 \omega_0 T_2}{\hbar} \quad (\text{A-8})$$

and

$$\gamma = \frac{A N_0}{\omega_0} \quad (\text{A-9})$$

Thus we see that the rate equations formalism and the formalism used in this report are equivalent when applied to individual lasers. Of course, the rate equations are just particle bookkeeping equations. They do not describe the electromagnetic wave and thus cannot be used in formulating the synchronization problem.

2.5-Bit/Detected Photon Demonstration Program: Description, Analysis, and Phase I Results

J R Lesh, J Katz, H H Tan, and D Zwillinger
Communications System Research Section

An optical communications demonstration program designed to demonstrate 2.5 bits of information transfer per single detected photon is described. The program was divided into four demonstration phases representing increasing degrees of complexity. The theoretical analysis of each of the phases is presented where it is shown that the 2.5-bit/detected photon goal can be achieved. The experimental results of phase I, which are in excellent agreement with the theory, are also presented.

I. Introduction

As deep space missions reach to greater distances and the scientific objectives of such missions become more ambitious the requirements for greater communication capability will continue to challenge the communications engineer. Past challenges have resulted in the transition from the S-band to the X-band frequency range, the increase in spacecraft antenna size and the arraying of multiple ground receive stations. Indeed the physical appearance of the Voyager-class spacecraft is completely dominated by the high-gain communications antenna. One cannot help wondering how much further such dimensions can be extended before reaching the limits of practicality.

One solution to this dilemma is to utilize optical frequencies for communications (Ref. 1). At optical frequencies the aperture determining components are usually measured in centimeters rather than meters. Furthermore, with the emergence of semiconductor injection lasers, one can reasonably anticipate that rugged, efficient solid-state optical sources will be available in the near future (Ref. 2). But even more significant is the fact that for receivers employing direct photon detection the communications channel capacity is extremely

large, with only a moderate amount of optical signal energy. This point was discussed by Pierce in 1978 (Ref. 3) where he described a channel capable of conveying tens of bits of information for each signal photon. This might seem quite surprising since the classical (nonquantum) theory predicted a limit of only one nat of information per photon ($1 \text{ nat} = 1.44 \text{ bits}$). However, Pierce was really illuminating and expanding upon predictions that had been made some time earlier (Refs. 4-7) and which have now stood the test of time.

Although channel capacity is recognized as the ultimate throughput limit of the channel, it leaves unanswered the question of how one can approach that limit in practice. For this reason a number of studies were initiated to find good modulation and coding schemes for the optical channel. It is now known that by properly matching these two areas, transmission of up to 5 bits/photon is achievable with today's technology (Refs. 8, 9). All of these schemes employ M -ary pulse position modulation and either Reed-Solomon codes or interleaved, short-constraint-length (multiple) convolutional codes. Furthermore, achieving energy efficiencies greater than 5 bits/photon is impeded only by the present-day limits of signal processing complexity, limits which are rapidly receding with the onslaught of VLSI technology.

Thus far only theoretical performance calculations have been carried out. The results of such calculations depend very heavily on the validity of the mathematical models used to describe the system. For this reason a demonstration program was initiated to prove out the validity of the models and with the prime objective of demonstrating reliable communications at a data rate of at least 100 kbps and an energy efficiency of 2.5 bits/detected photon. (Reliable communications was defined as the Voyager threshold bit error rate of 5×10^{-3} .) The 2.5-bit/photon number was selected because it is comfortably above the classical limit but sufficiently small so that the system complexity would not get out of hand. After flushing out the problems at this design point a more ambitious program to demonstrate 5 or more bits/detected photon will be initiated.

In the next section we will describe the 2.5-bit/detected photon demonstration program and show how it has been divided into four program phases. In Section III we will present a theoretical analysis of the system and show that the primary demonstration goal can theoretically be achieved. This analysis is substantially more detailed than previous studies which tended to ignore such things as photomultiplier tube gain dispersion and thermal noise in the predetection pulse amplifier. In Section IV we will then describe the experimental results of the first phase of the program which show that the detector noise models used in the analysis are comfortably close to reality. Finally, the results are summarized and discussed in Section V.

II. General Description of the 2.5 Bit/Detected Photon Program

A block diagram of the demonstration system is shown in Fig. 1. The heart of the optical portion of the system consists of a gallium arsenide semiconductor injection laser and a direct detection photomultiplier tube. The GaAs laser diode is a high-quality, single-spatial-mode device operating at $0.85 \mu\text{m}$ and has the reliability and durability characteristic of solid-state devices. The light emitted from the laser is passed through some elementary optics followed by 70-100 dB of neutral density filters (attenuators) which simulate space loss. The attenuated optical signal is then applied to a photomultiplier tube (PMT) detector which has a high internal gain ($>10^6$), a quantum efficiency around 20% and, with moderate cooling, an extremely low dark current. Of course, to eliminate stray laboratory light, all of the optical components must be placed in an extremely dark enclosure.

Surrounding the optical components are the modulation and coding hardware. The laser diode is driven by a 256-slot/word PPM modulator which decides, based on an 8-bit input

word, which slot the pulse should be placed in and then provides a current pulse during that slot to turn on the laser. Obviously, the inverse of this process is applied to the PMT output to recreate the 8-bit word. To improve the performance of the system, an 8-bit Reed-Solomon code is then used to surround the PPM portion of the system. An 8-bit Reed-Solomon encoder considers 8-bit segments of the incoming (binary) data stream as individual (generalized) symbols and then performs error correction coding over the strings of these generalized symbols. Since the code symbol size and PPM word size are matched, PPM word erasures or errors correspond to a single Reed-Solomon code symbol erasures or errors respectively. Such codes are well known for their burst erasure fill-in capabilities and, furthermore, can compensate for combinations of errors and erasures. The data streams supplied to the encoder and delivered from the decoder are then compared for an overall bit error rate measurement.

The demonstration program has been divided into four phases. The first phase involves only the PMT and its associated preamplifier and is intended to characterize the dark current noise distribution of the detection system. This phase has essentially been completed and the results will be described in Section IV. In the second phase (currently in progress) the laser will be added and the optical pulse erasure and error statistics will be evaluated. The PPM modulator and demodulator, which have already been designed and fabricated, will be added in phase III so that the PPM word error and word erasure probabilities can be measured. The final phase will encompass the coding hardware and will demonstrate the 2.5-bit performance goal.

III. Theoretical Analysis of the Demonstration

In this section we will analyze the performance of the 2.5-bit/detected photon demonstration system. A theoretical model of a photomultiplier tube and associated preamplifier was developed in Ref. 10. This model includes the effects of the photomultiplier tube random gain as well as post PMT preamplifier noise and is useful for computing communication system error performance as well as estimating, from experimental data, system parameters such as the PMT dark current. We will first describe the results in Ref. 10 pertaining to the model of the PMT gain random variable and the calculation of its probability density function. Then, using this model we will determine the detection and false alarm probabilities of a photocathode electron emission event. These results will be useful in characterizing both the dark current photoelectron emission event statistics (phase I) as well as optical pulse detection statistics (phase II). Next the pulse detection statistics are used to evaluate the error performance of the PPM

(uncoded) signal set. Finally, the coded PPM performance will be analyzed.

A. Random Photomultiplier Tube Gain

In a photomultiplier tube, electrons emitted by the photocathode (either from dark current or photon absorption) are directed through a series of dynodes by an electric field. Secondary electrons are emitted at each dynode for each impinging primary electron. The number of secondary electrons at each dynode is a random quantity and results in a random overall gain for the PMT. The mean and variance of this gain can be readily derived. An explicit expression for the exact probability distribution of the gain, however, is not available and the problem of determining it appears to be intractable. Since an analytical expression for this distribution is required to determine communication system error performance, a good approximation of the exact distribution is useful. In Ref. 10, such an approximation was derived by using a diffusion model. The approximate density $p_G(x)$ of the random gain G is given by

$$P_G(x) = e^{-\bar{G}/B} \left\{ \delta(x) + \left(\frac{1}{B} \sqrt{\frac{\bar{G}}{x}} \right) e^{-x/B} I_1 \left(2 \frac{1}{B} \sqrt{\bar{G}x} \right) \right\}, \quad (1)$$

where $\delta(x)$ is the unit impulse function, $I_1(\cdot)$ is the modified Bessel function of the first kind,

\bar{G} = average gain of the photomultiplier tube

ν = number of dynode stages

and where the parameter B is specified by \bar{G} and ν according to

$$B = \frac{1}{2} \frac{\bar{G} - 1}{\sqrt{\bar{G}} - 1} \quad (2)$$

This approximate density yields the correct mean and variance for G . In fact, the variance of G is simply given by $2\bar{G}B$. For the RCA C31034 PMT used in this demonstration \bar{G} is about 10^6 and $\nu = 11$, so the standard deviation of G is about $0.6\bar{G}$ and B is equal to about one-third of the standard deviation. Figures 2(a) and 2(b) plot the density of the normalized gain G/\bar{G} as given by (1) for the PMT used in the demonstration at $\bar{G} = 10^6$ and 10^7 respectively. As can be seen from these figures, the density is not symmetric about the mean and has substantial probability mass in the region below its mean value \bar{G} . This means that there is substantial likelihood that the photomultiplier output signal level will be lower than the

corresponding average gain value, and error performance calculations based on a constant gain model will not be accurate. Also shown in Fig. 2 is a density function that is the positive truncation of the Gaussian density having the same mean and variance as $p_G(x)$. It is clear that using the simpler Gaussian approximation can produce results that are substantially optimistic.

B. Photocathode Emission Event Detection: Receiver Operating Curves

In phase I of the demonstration the number of dark current emission events at the PMT cathode is experimentally recorded by monitoring the PMT output with a pulse counter. In Ref. 10, the effects of random PMT gain, thermal noise in the post PMT preamplifier, and counter response time on successful photocathode dark current emission event detection were evaluated. The PMT anode output is assumed to be terminated by a resistor of R ohms. The voltage signal $S(t)$ across this resistor is amplified and fed into a pulse counter. The amplifier-counter combination is modeled as an additive thermal noise $V(t)$ followed by a short term time averager over T_c seconds (the counter response time). The output of this time averager is then compared with a threshold γ . An emission event is recorded by the counter only when the threshold γ is exceeded. The entire system is shown in Fig. 3. The additive thermal noise $V(t)$ is assumed to have spectral density

$$N_0 = k\theta R \text{ (volts)}^2/\text{Hz} \quad (3)$$

where k = Boltzmann's constant, θ = noise equivalent temperature of the amplifier and R = amplifier equivalent resistance (assumed to be matched to the anode terminating resistance). If P_d = detection probability of an emission event and P_f = false alarm probability due to thermal noise, then it is shown in Ref. 10 that

$$P_d = \int_0^\infty Q\left(\frac{\gamma - \left(\frac{eR}{T_c}\right)x}{\sigma}\right) P_G(x) dx \quad (4)$$

and

$$P_f = Q\left(\frac{\gamma}{\sigma}\right), \quad (5)$$

where $p_G(x)$ is given by (1) and where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt,$$

and

$$\sigma^2 = \frac{N_0}{T_c}$$

Receiver operating curves of P_d versus P_f are shown in Fig 4 for a counter response time $T_c = 10$ ns, thermal noise standard deviation $\sigma = 34 \mu\text{V}$ and for several values of PMT average gain \bar{G} . Also shown on the figure are the corresponding threshold values

These curves are useful in interpreting pulse counter data for the purpose of determining whether the number of dark current photoelectrons is Poisson distributed. Suppose pulse counter data is recorded over a period of 1 sec and the counter response time $T_c = 10$ nsec. Assume that the average gain of the photomultiplier $\bar{G} = 10^6$ and the PMT is cooled to -30°C . With these parameters the RCA C31034 PMT is rated to have an average dark current emission rate at the photocathode of about 25 counts per second. Then under an assumption of Poisson statistics, the probability of one dark current photoelectron emission event in an interval of $T_c = 10$ nsec is about

$$\lambda_n = 25 T_c = 2.5 \times 10^{-7}$$

and the probability of no dark current events is approximately $1 - \lambda_n$. The probability that the pulse counter records one count in a T_c sec slot time is then given by

$$p = \lambda_n P_d + (1 - \lambda_n) P_f$$

and the probability of no counts is equal to $1 - p$. In a 1-sec observation interval there are $n = 1/T_c = 10^8$ T_c -sec slots. Thus, under a Poisson assumption on the dark current photoelectron emission statistics, the number of pulse counter counts in a 1-sec observation period is a binomial (n, p) random variable.

Suppose now that the threshold γ is set at $200 \mu\text{V}$, $\sigma = 34 \mu\text{V}$ and $\bar{G} = 10^6$. Then the receiver operating curve gives $P_d = 0.91$ and $P_f = 2 \times 10^{-9}$. Under these conditions $np \approx n\lambda_n P_d = 23$ when $n = 10^8 \gg 1$. In this case the binomial (n, p) distribution is approximately Poisson with intensity np , so the number of recorded pulse counter counts will still be Poisson even with the random photomultiplier gain and the amplifier thermal noise taken into account. Next, suppose the threshold γ is set lower at $100 \mu\text{V}$. Here the receiver operating curve gives $P_d = 0.962$ and $P_f = 2 \times 10^{-3}$. In this case $np \approx nP_f = 2 \times 10^5 \gg 1$ when $n = 10^8$ and the binomial (n, p) distribution is approximately Gaussian with mean np and variance $np(1 - p)$. The pulse counter data will no longer be

Poisson. Therefore, care must be taken in using pulse counter data for determining counting statistics. Specifically, if the counter threshold is set too low, the photocathode emission false alarm rate will be too high and the counter statistics will be dominated by the thermal (Gaussian) noise of the PMT preamp.

C. Noise and Signal Pulse Detection: Receiver Operating Curves

Whereas phase I of this demonstration was concerned with the dark current statistics of the PMT, in phase II the laser is added to evaluate the error performance of the PMT-preamplifier combination in detecting the presence of laser light excitation. In this demonstration phase, the laser is either on or off in a time slot of duration T_s seconds. The PMT amplifier output is integrated over this time period, normalized by T_s and compared to a threshold γ . Detection of the laser light is declared if and only if the threshold is exceeded. In Ref. 10, the error performance of this detector was evaluated assuming that photocathode emission events are Poisson distributed and taking into account the random gain density (1) as well as the preamplifier thermal noise. Specifically, consider the detector shown in Fig 5, where the amplifier is assumed to introduce an additive white Gaussian noise $V(t)$ with spectral density N_0 given by (3) to the signal $S(t)$ across the anode resistor R . Let

$$P_{ds} = \text{probability of correctly detecting the presence of incident laser light} \quad (6)$$

and

$$P_{dn} = \text{probability of correctly detecting the absence of incident laser light} \quad (7)$$

These detection probabilities P_{ds} and P_{dn} depend on the thermal noise level N_0 , the threshold γ , the integration time T_s , and the average gain \bar{G} and dark current of the PMT, as well as the incident light intensity at the photocathode. Specifically let

$$\bar{N}_n = \text{average number of dark current photocathode emissions per } T_s \text{ sec time period at the photocathode} \quad (8)$$

and

$$\begin{aligned} \bar{N}_s &= \text{average number of light-excited photocathode emission events per } T_s \text{ sec time period} \\ &= \frac{\eta P_s}{hf}, \end{aligned} \quad (9)$$

where η is the photocathode quantum efficiency, P_S is the incident signal light intensity at the photocathode surface, h is Planck's constant and f the incident light center frequency. Also define

$$\sigma_1^2 = N_0/T_s$$

and let

$$P(\alpha) = e^{\alpha(e^{-\bar{G}/B} - 1)} Q\left(\frac{\gamma}{\sigma_1}\right) + \frac{\sqrt{G}}{B} e^{-\alpha} \sum_{n=0}^{\infty} \frac{\sqrt{n}(\alpha e^{-\bar{G}/B})^n}{n!} J_n \quad (10)$$

where

$$J_n = \int_0^{\infty} Q\left(\frac{\gamma - \left(\frac{eR}{T_s}\right)y}{\sigma_1}\right) \frac{e^{-y/B}}{\sqrt{y}} I_1\left(2 \frac{\sqrt{nGy}}{B}\right) dy$$

and B is given by (2). It was shown in Ref. 10 that

$$P_{ds} = P(\alpha) \Big|_{\alpha = \bar{N}_s + \bar{N}_n}, \quad (11)$$

and

$$1 - P_{dn} = P(\alpha) \Big|_{\alpha = \bar{N}_n} \quad (12)$$

Typical on-off pulse detection receiver operating curves of P_{ds} versus $1 - P_{dn}$ are shown in Fig. 6 for integration time $T_s = 100$ ns, thermal noise standard deviation $\sigma_1 = 34$ μ V and for several values of \bar{N}_s . Figure 6(a) shows the results for $\bar{G} = 10^6$ and 25 dark current counts/second. Figure 6(b) shows the corresponding results for $\bar{G} = 3.2 \times 10^6$ and 53 dark current counts/sec¹

¹The PMT average gain can be increased by increasing the relative anode voltage. Gains from 10^5 to 10^7 can be produced in this way. However, increasing the anode voltage results in an increase in the number of dark current electrons drawn from the photocathode. The values of dark current used in Fig. 6 correspond to the RCA 31034 PMT operating at -30°C at the various values of gain (see Fig. 12 of Section IV).

D. Uncoded PPM Performance

Phase III of the demonstration involves the addition of pulse position modulation and demodulation equipment to the optical and elementary electronic signal detection equipment. In M -ary PPM signaling M adjacent time slots, each of duration T_s sec, are used to form a single PPM symbol. A given T_s second slot is either a "noise slot," in which case no light is emitted from the laser diode transmitter, or a "signal slot" when light of a constant intensity is emitted from the laser. In a PPM symbol time one, and only one, of the slots is a signal slot and the remaining $(M - 1)$ slots are noise slots. Since the information to be transmitted is contained in the position (in time) of the signal slot, each PPM symbol can convey $\log_2 M$ bits of information (if perfectly detected). In the 2.5-bit/detected photon demonstration $M = 256$, but we shall perform the analysis for arbitrary M .

The PPM demodulator receives signal slot/noise slot decisions over the slot time T_s from the PMT and light detector circuit of Fig. 5. Since separate decisions are provided to the demodulator for each of the M slots in a PPM symbol, a number of different types of error events are possible. We will analyze each of these event statistics below.

1. PPM symbol erasure probability. When no signal slot decisions are received during an entire PPM symbol a symbol erasure is declared. The probability of this event, P_E , is the probability that the signal slot is not detected and that the remaining $(M - 1)$ noise slots are correctly classified. In the notation of Fig. 6,

$$P_E = (1 - P_{ds}) P_{dn}^{M-1} \quad (13)$$

which can be easily evaluated given an operating point on that figure.

2. Undetected PPM symbol error probability. An undetected symbol error occurs when the signal slot is incorrectly declared a noise slot and one (and only one) of the noise slots is taken as a signal slot. Since the incorrectly detected noise slot can occur at any one of $(M - 1)$ places, the probability, P_{us} , of an undetected symbol error is

$$P_{us} = (M - 1) (1 - P_{ds}) (1 - P_{dn}) P_{dn}^{M-2} \quad (14)$$

3. Correct PPM symbol detection probability. When a PPM symbol is transmitted it will be received correctly if the signal slot and all of the noise slots are correctly detected. The probability of this event is

$$P_c = P_{ds} P_{dn}^{M-1} \quad (15)$$

4 Two or more pulse probability. When two or more pulses are received during a PPM symbol time the demodulator has conflicting instructions. The probability of this event, P_{2P} is easily derived as follows

$$\begin{aligned} P_{2P} &= Pr(\text{no pulses detected}) - Pr(1 \text{ pulse detected}) \\ &= 1 - P_E - P_c - P_{us} \\ &= 1 - P_{dn}^{M-1} - (M-1)(1 - P_{ds})(1 - P_{dn})P_{dn}^{M-2} \quad (16) \end{aligned}$$

5 Uncoded PPM symbol error probability. The above events are all measurable during the phase III tests and are important in determining the overall performance of the system when coding is later applied. However, it is also possible to consider the system as purely an uncoded PPM system during phase III. If this is to be accomplished then some resolution of erasures and multiple pulses per symbol time must be accomplished. A maximum likelihood demodulator would use the following decision rules

- (1) If all M slots are declared noise slots, choose the transmitted signal slot at random from the M slots
- (2) If exactly $k \geq 1$ slots are declared signal slots, choose the transmitted signal slot at random from among those k slots

The actual PPM demodulator does not use these rules. Whenever an erasure occurs it always selects a preassigned symbol as its output. Also, if two or more pulses are received in a symbol time, the output symbol corresponds to the first detected signal slot. However, since the transmitted PPM symbols are selected at random and the channel is assumed to be memoryless, the performance of the actual decoder will be the same as the maximum likelihood version. We will therefore analyze the latter.

Since the demodulator always produces an output symbol then the probability of a symbol error, P_S , is 1 minus the probability of correct symbol detection. (Note that the correct symbol probability is *not* P_c as developed above. This is because P_c assumes that neither erasures nor multiple pulses are resolved. P_c is therefore only one component in the uncoded PPM probability of correct symbol detection.) We have, therefore, that

$$\begin{aligned} P_S &= 1 - P_r(\text{correct symbol detection}) \\ &= 1 - P_{ds}(P_{dn})^{M-1} - \frac{1}{M}(1 - P_{ds})P_{dn}^{M-1} \end{aligned}$$

$$- \sum_{k=1}^{M-1} \frac{1}{k+1} \binom{M-1}{k} P_{dn}^{M-1-k} (1 - P_{dn})^k P_{ds} \quad (17)$$

where the second term corresponds to P_c , the third term is the probability of correct decision when an erasure occurs and the last term corresponds to the probability of a correct symbol selection when the actual signal slot and one or more noise slots are declared signal slots. In the appendix it is shown that (17) simplifies to

$$P_S = 1 - \frac{1}{M}(1 - P_{ds})P_{dn}^{M-1} - \frac{P_{ds}(1 - P_{dn}^M)}{M(1 - P_{dn})} \quad (18)$$

6. Uncoded PPM bit error probability. The final quantity of interest is the probability of a bit error for the uncoded PPM configuration. Since PPM is a form of orthogonal modulation, the bit error probability P_b and the PPM symbol error probability can be related by

$$P_b = \left(\frac{\frac{M}{2}}{M-1} \right) P_S \quad (19)$$

During the phase III evaluation P_b will be one of the key performance parameters. Figure 7 shows P_b as a function of the information power efficiency measure $\rho = \log_2 M / N_s$ (bits/photon) for several values of the analog detector threshold γ , $M = 256$, $\sigma = 34 \mu V$ and $T_s = 100$ ns. Figure 7a shows the result for $\bar{G} = 10^6$ and 25 dark current counts/second whereas Fig 7b is for $\bar{G} = 3.2 \times 10^6$ and 53 counts/second. In both figures the performance improves as the threshold γ is reduced from 250 to 150 μV . This is because the pulse erasure probability is also decreasing. However, if the threshold is reduced too far, the noise generated false alarm rate increases, which in turn increases P_b . This effect is evident in the $\gamma = 100 \mu V$ curves. Nevertheless, even with an optimization of the threshold it is clear that without the benefits of the overall coding system, operation at a reasonable error rate is not possible at 2.5 bits/photon.

E. Coded PPM Performance

In the final demonstration phase, a Reed-Solomon (R-S) encoder and decoder are used over the PPM modulated channel. Since the code symbol alphabet size is 8 bits (matched to the PPM symbol size), the (R-S) code blocklength, N , is $2^8 - 1 = 255$ (symbols) (Ref. 11). Two specific codes were considered for the demonstration. The first is the (255,223), $t = 16$ error correcting, rate 7/8 (approx.) code. The second is the (255,191), $t = 32$ error correcting code which has a rate of approximately 3/4.

A t error correcting (R-S) code can correct any combination of t or less symbol errors. It can also compensate for combinations of errors and erasures. In particular, if s is the number of decoder input symbol errors and e is the number of decoder input symbol erasures, then the decoder will produce the correct output codeword provided

$$2s + e < d_{\min} \triangleq 2t + 1 \quad (20)$$

The code is therefore more powerful at correcting erasures than errors

One can easily show that if the confidence that a particular decoder input symbol is correct is strictly less than 0.5, it is better to erase the symbol and allow the decoder to attempt to fill it in. This condition is always satisfied whenever a PPM erasure occurs or whenever two or more pulses are received in a PPM symbol time (The case of two pulses can also result when the signal slot is not detected and two noise slots are declared signal slots). Thus, when (R-S) coding is used, it is necessary to modify the demodulator decision rules. Toward this end let us define a decoder input symbol error event to be the occurrence of an undetected PPM symbol error, and a decoder input symbol erasure as either a PPM symbol erasure or the occurrence of two or more pulses in a PPM symbol time. Then, from Eqs (13 – 16), the probabilities of decoder input symbol error $P_{\mathcal{S}}$ and symbol erasure $P_{\mathcal{E}}$ are given by

$$P_{\mathcal{S}} = P_{us} = (M - 1) (1 - P_{ds}) (1 - P_{dn}) P_{dn}^{M-2} \quad (21)$$

and

$$\begin{aligned} P_{\mathcal{E}} &= P_E + P_{2P} = 1 - P_{ds} P_{dn}^{M-1} \\ &- (M - 1) (1 - P_{ds}) (1 - P_{dn}) P_{dn}^{M-2} = 1 - P_c - P_{us} \end{aligned} \quad (22)$$

Once we have $P_{\mathcal{S}}$ and $P_{\mathcal{E}}$ we can compute the decoder output performance. A (R-S) decoder will always produce the correct codeword at the output whenever the error/erasure capability of the code (Eq. 20) is satisfied. When this condition is not satisfied the decoder finds that it cannot solve the equations which determine the location of the codeword symbol errors nor can it solve for the values of the erased symbols. Consequently, a decoding (computational) failure results. The codes considered are systematic, which means that each codeword is made up of $255-2t$ encoder input information symbols followed by $2t$ generalized parity symbols (which are computed in the encoder). When a decoding failure occurs, the decoder will output the systematic portion of the codeword.

This means that for that particular codeword the power of the code is not (actually cannot be) used, but the uncoded portion of the codeword is still more likely to be correct than a randomly selected codeword

The probability that the incorrect codeword is selected by the decoder, P_{WE} , can be written by inspection as

$$P_{WE} = \sum_{s=0}^N \sum_{e=\Delta}^{N-s} \binom{N}{s} \binom{N-s}{e} P_{\mathcal{S}}^s P_{\mathcal{E}}^e (1 - P_{\mathcal{S}} - P_{\mathcal{E}})^{N-s-e} \quad (23)$$

where

$$\Delta = \max(d_{\min} - 2s, 0).$$

The argument of the summation in (23) is recognized as the probability that a received codeword contains s symbol errors, e erasures and the remaining $N-s-e$ symbols are error- and erasure-free. The summation is then taken over all values of s and e which violate Eq. (20). The symbol $=$ is used since (23) is actually a slight overbound. This is because there is some chance (albeit extremely small) that, even though (20) is violated, the systematic portion of the codeword is received without error.

The probability of a decoder output symbol error P_{SE} can now be computed by assuming that when s and e errors and erasures respectively occur in a codeword, they are statistically spread uniformly over the N codeword symbols. Then assuming that all symbol erasures are (pessimistically) resolved in error at the decoder output gives

$$\begin{aligned} P_{SE} &= \sum_{s=0}^N \sum_{e=\Delta}^{N-s} \binom{N}{s} \binom{N-s}{e} \\ &\cdot P_{\mathcal{S}}^s P_{\mathcal{E}}^e (1 - P_{\mathcal{S}} - P_{\mathcal{E}})^{N-s-e} \left(\frac{s+e}{N} \right) \end{aligned} \quad (24)$$

Given P_{SE} we can also compute P_{BE} , the decoder output bit error rate from the expression

$$P_{BE} = \frac{M}{2(M-1)} P_{SE} \quad (25)$$

Equation (25) is the final performance measure for the 2.5-bit/detected photon demonstration. Figure 8 shows P_{BE} as a function of the energy efficiency parameter ρ for a PMT gain

of 3.2×10^6 , 53 dark counts/second, $T_s = 100$ ns and several values of pulse detector threshold voltage γ . Figure 8(a) illustrates the performance of the rate 7/8 code whereas Fig 8(b) applies to the rate 3/4 code. It is clear from these figures that the rate 7/8 code will not produce an adequate bit error rate at $\rho = 2.5$ bits/photon. The rate 3/4 code, on the other hand, provides Voyager threshold performance up to 2.9 bits/photon for $\gamma = 150 \mu\text{V}$.

The corresponding performance curves for a PMT gain of 10^6 were not computed since it was recognized that they would not provide adequate performance. This can be seen from the following analysis. From Fig 7(a) we see that the uncoded PPM bit error rate at $\rho = 2.5$ bits/photon and a PMT gain of 10^6 is greater than 0.15 for all values of γ . From (19) this corresponds to a PPM symbol error rate of 0.3, which results from undetected symbol errors and incorrectly resolved ambiguities. This implies that on the average a (R-S) codeword of length 255 will have 76 anomalies, either symbol errors or erasures. Since this number exceeds the code correction capabilities (Eq. 20) of both codes *even if all anomalies are erasures*, the coded performance will obviously be inadequate. Contrast this with the 3.2×10^6 gain results of Fig 7(b). At $\rho = 2.5$ and $\gamma = 150 \mu\text{V}$ the average number of anomalies per codeword is 20. This satisfies Eq. (20) for the rate 7/8 code only if most are erasures. For the rate 3/4 code, however, (20) is satisfied on the average regardless of whether the anomalies are erasures or errors.

One additional comment is in order. The demonstration goal is to achieve 2.5-bit/detected photon performance at 100 kbps. The performance calculations in this section were made for $T_s = 100$ ns, which corresponds to the maximum clock rate of the logic used in the hardware. For 256-ary PPM and rate 3/4 coding, these results correspond to 234 kbps.

IV. Experimental Results of Phase I

A. Experimental Setup and Conditions

As described earlier, the purpose of this phase of the work was to measure the intrinsic noise characteristics of the system and to verify the theory used to model it.

The block diagram of the experimental setup is shown in Fig. 9. The photomultiplier tube (PMT) used in the experiment was the RCA model C31034. This PMT has a GaAs photocathode with the highest quantum efficiency available at the wavelength region where AlGaAs lasers operate (0.8–0.9 μm). This particular PMT type has a quantum efficiency (η) of about 15% at 0.85 μm . A selected type of this PMT – Model C31034A-02 – with $\eta \cong 20\%$ was also purchased but was not used in the experiment. The above efficiency values

are at room temperature, and a moderate improvement is obtained with some cooling. The typical gain of this PMT is 10^6 . The gain varies with varying experimental conditions, as will be described below.

Since the dark current of the PMT is too high at room temperature, the PMT was placed in a cooled housing. The cooling reduces the thermionic emission from the PMT's photocathode, which is a dominant noise mechanism. The thermoelectric (T.E.) cooler used was a Model TE-206TS-RF (Products for Research, Inc.), which can cool the photocathode to a temperature 60 K lower than room temperature. Due to other noise mechanisms (Ref. 12), some temperature-independent residual dark current was observed below about -30°C . The high voltage (1200–1800V) is supplied to the PMT by a Model 204-10 power supply (Pacific Photometric Instruments).

The output of the PMT was directly connected to a fast preamplifier Model 9301 (ORTEC). This amplifier has a gain of 10, bandwidth of 150 MHz and equivalent input noise voltage of about 34 μV rms.

The PMT and preamplifier (which must be in close proximity because of noise considerations) were placed in a dark enclosure to eliminate ambient illumination and permit the observation of the actual PMT and preamp noise mechanisms. The effectiveness of the enclosure was verified by noting that no change in the dark count reading resulted when the laboratory was darkened or fully illuminated. Following the preamplifier was a second amplifier (HP Model 461) which provided final pulse gain (gain X100, bandwidth 150 MHz).

The amplified signal was first fed to a fast storage oscilloscope (Tektronix 7834) to verify that the PMT noise output was generated by single photocathode emission events, not bursts of events (Ref. 13). In all subsequent tests the amplified signal was connected to a HP5370A universal counter, which in turn was connected to a HP9845C computing controller through an HP-IB interface.

In the course of the measurements there were basically three parameters that could be changed. The temperature of the photocathode, the PMT voltage and the threshold (or trigger) level of the counter. Changing the temperature alters the intensity of the noise process, i.e., the number of photoelectrons released from the photocathode via thermionic emission (Ref. 14). Changing the PMT voltage V changes the average gain \bar{G} of the PMT according to

$$\bar{G} = aV^b$$

where V is in kilovolts, \bar{G} is in units of 10^6 and where $a \approx 0.043$ and $b \approx 9.22$ are constants calculated from the manufacturer's data sheet. Finally, changing the threshold level changes both the intensity of the dark current counts and their origin. At low threshold voltages most of the counts are due to white (Gaussian) thermal noise of the preamplifier, at higher values of threshold voltages the dominant noise mechanism is due to the dark counts originating at the PMT.

Several limitations on the measurements are imposed due to hardware constraints. First, the bandwidth of the preamplifier-amplifier combination is slightly less than 110 MHz, this matches the 100-MHz capability of the counter, but is somewhat less than the 140-MHz intrinsic bandwidth of the PMT. Secondly, when the counter operates in the time interval mode (as needed for the time-of-arrival distribution measurement), it counts only every other interval and, in addition, has a 330- μ s deadtime after each measurement. This limitation is not significant if we assume that the process is stationary and when operating in the Poisson regime, but for low values of threshold voltage, only the frequency mode of the counter can be used in order to obtain reliable results.

B. Experimental Results

The first experiment set was designed to test the validity of the assumption that the dark count events are generated by a Poisson process. To eliminate counts due to preamplifier thermal noise the counter threshold voltage was set at the 9σ level of the thermal process. A sample set of 1000 count events was then observed and compared in two ways against theoretical predictions. The first comparison involved the computation of the number of events in a preset time interval to verify that it was Poisson distributed. The second involved computation of the event interarrival times to determine if they were exponentially distributed. Typical experimental results are shown in Fig. 10 along with the associated theoretical distributions. Figure 10(a) shows the frequency distribution (variation in the number of counts per time interval), whereas Fig. 10(b) contains the interarrival time distributions. The results show extremely good agreement between theory and experiment. Additionally, two χ -square tests appropriate for the comparisons of these distributions were computed. These tests confirmed the visual conclusions that the theoretical and experimental distributions were close. Similar results were obtained for several different values of PMT gains.

Measurements were also made at lower values of the counter threshold voltage. Here, as expected, the measured distributions deviated considerably from the Poisson model due to the increased number of counts resulting from preamplifier thermal noise.

The second experiment set involved taking extensive data to determine the average intensity of the dark current Poisson process, as well as its variation with system parameters. First, the dependence of the dark current intensity on PMT temperature was measured and compared with the manufacturer's specifications. Favorable agreement was obtained which provided confidence that the experimental setup was performing correctly. Then, the effects of PMT gain and counter threshold voltage on the dark current intensity were measured. To obtain reliable results, 10^4 seconds of data were collected for each gain and threshold setting. The results are shown in Fig. 11. As expected, the average intensity of noise events decreases with an increase in the counter threshold voltage.

Another important feature of Fig. 11 is the rather large dependence of the dark current intensity on the PMT gain setting. This phenomenon is due to the fact that as the voltage of each dynode is increased, the PMT enters an unstable region produced by regenerative ionization effects (Ref. 12). It is important that this variation be adequately modeled so that optimized systems designs can be obtained. However, the data of Fig. 11 includes not only the effects of the PMT but the preamplifier and counter threshold as well. By modeling the overall process and working backward from the experimental results one can deduce the intrinsic variation in PMT dark current intensity as a function of PMT gain. This result is shown in Fig. 12. The figure clearly indicates an exponential dark current dependence on average gain.

V. Conclusions

We have described the 2.5-bit/detected photon demonstration program and identified how it has been broken into four demonstration phases. Then an analysis of each phase was presented. The analysis illustrated how, with 256-ary PPM modulation, Reed-Solomon coding and carefully chosen system parameters, it is possible to achieve the 2.5-bit/detected photon goal at rates above 100 kbps. Finally, the results of the phase I experiments were presented where the agreement between theory and experiment was shown to be surprisingly close. Additionally, the parametric dependencies of the PMT detection noise intensities were also experimentally determined.

The agreement in the phase I tests provides a substantial amount of credibility to the analysis and associated conclusions presented herein. However, the final proof will come with the actual phase IV demonstration. Nevertheless, the results of phase I will permit the phase II-IV activities to be conducted with increased optimism.

Acknowledgment

The authors wish to express their appreciation to Gary Lorden for his assistance on the statistical data analysis algorithms used in Phase I

References

1. Katz, J , and Lesh, J R., "Optical Communications for Deep Space Missions Using Integrated Semiconductor Injection Lasers," Jet Propulsion Laboratory, Pasadena, Calif., (in preparation).
2. Katz, J., "High Power Semiconductor Lasers for Deep Space Communications," in *The TDA Progress Report 42-63*, pp. 40-50, Jet Propulsion Laboratory, Pasadena, Calif , June 15, 1981
3. Pierce, J R , "Optical Channels Practical Limits with Photon Counting," *IEEE Trans. Comm.*, Vol COM-26, No 12, pp 1819-1821, Dec 1978
4. Gordon, J. P , "Quantum Effects in Communications Systems," Proceedings of the Institute of Radio Engineers," Vol 50, pp 1898-1908, Sept. 1962.
5. Levitin, L. B , "Photon Channels with Small Occupation Numbers," *Problemy Predachi Informatsi*, Vol. 2, No. 2, pp 60-68, 1966
6. Helstrom, C. W., Liu, J W. S , and Gordon, J. P , "Quantum Mechanical Communication Theory," *Proc IEEE*, Vol 58, pp. 1578-1598, Oct 1970.
7. Helstrom, C. W., "Capacity of the Pure State Quantum Channel," *Proc IEEE*, Vol 62, pp. 140-141, Jan 1974.
8. McEliece, R. J., "Practical Codes for Photon Communication," *IEEE Trans Information Th.*, Vol IT-27, 1981.
9. Massey, J L , "Capacity, Cutoff Rate, and Coding for a Direct Detection Optical Channel," in *TDA Progress Report 42-60*, pp 68-76, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 1981.
10. Tan, H. H , "A Statistical Model of the Photomultiplier Gain Process with Applications to Optical Pulse Detection," Jet Propulsion Laboratory, Pasadena, Calif., (in preparation)
11. McEliece, R. J., *The Theory of Information and Coding*, Addison-Wesley, Reading, MA, 1977
12. Engstrom, R. W., "Photomultiplier Handbook," Ch 4, RCA Solid-State Div., Lancaster, PA 17604, 1980
13. McHose, R. E., "Time Characteristics of Photomultipliers – some General Observations," RCA application note AN-4884, 1972.
14. Katz, J , "The Deep Space Optical Channel I. Noise Mechanisms," *TDA Progress Report 42-64*, pp. 180-186, Jet Propulsion Laboratory, Pasadena, Calif , Aug. 15, 1981.

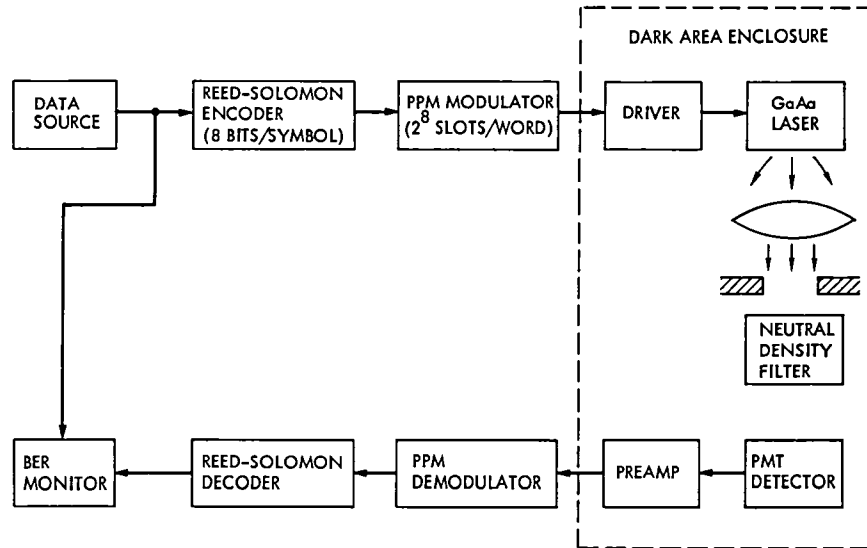


Fig 1 Block diagram of 2.5-bit/detected photon demonstration system

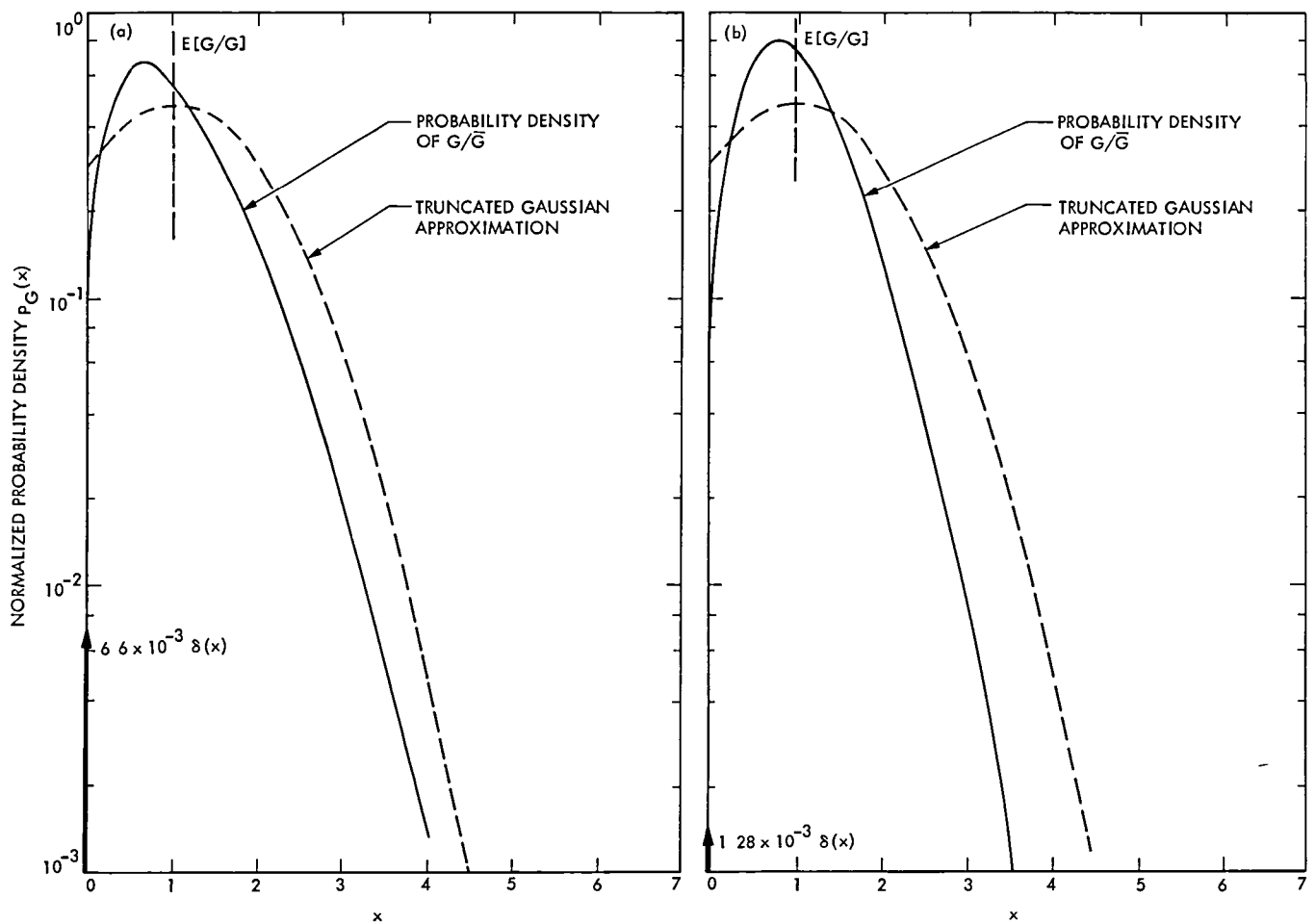


Fig. 2. Probability density function of the normalized gain of an eleven-stage PMT and (a) average gain = 10^6 , (b) average gain = 10^7

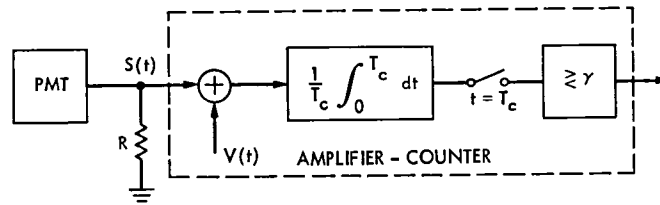


Fig. 3. Equivalent block diagram of the PMT noise model measurement system

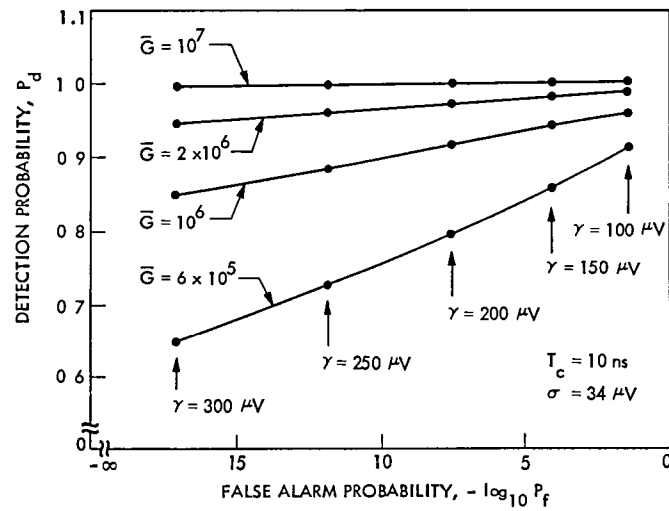


Fig. 4. Detection and false alarm probabilities (conditioned on a photocathode emission event)

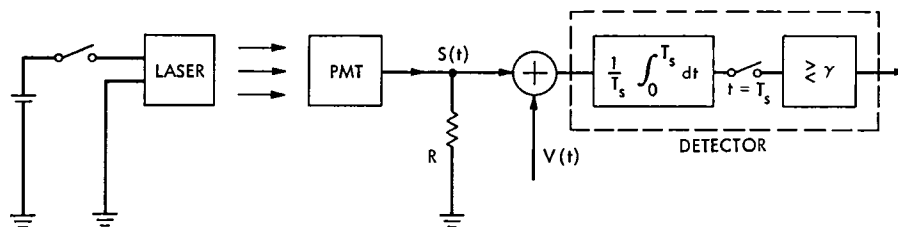


Fig. 5. Equivalent block diagram of phase II demonstration

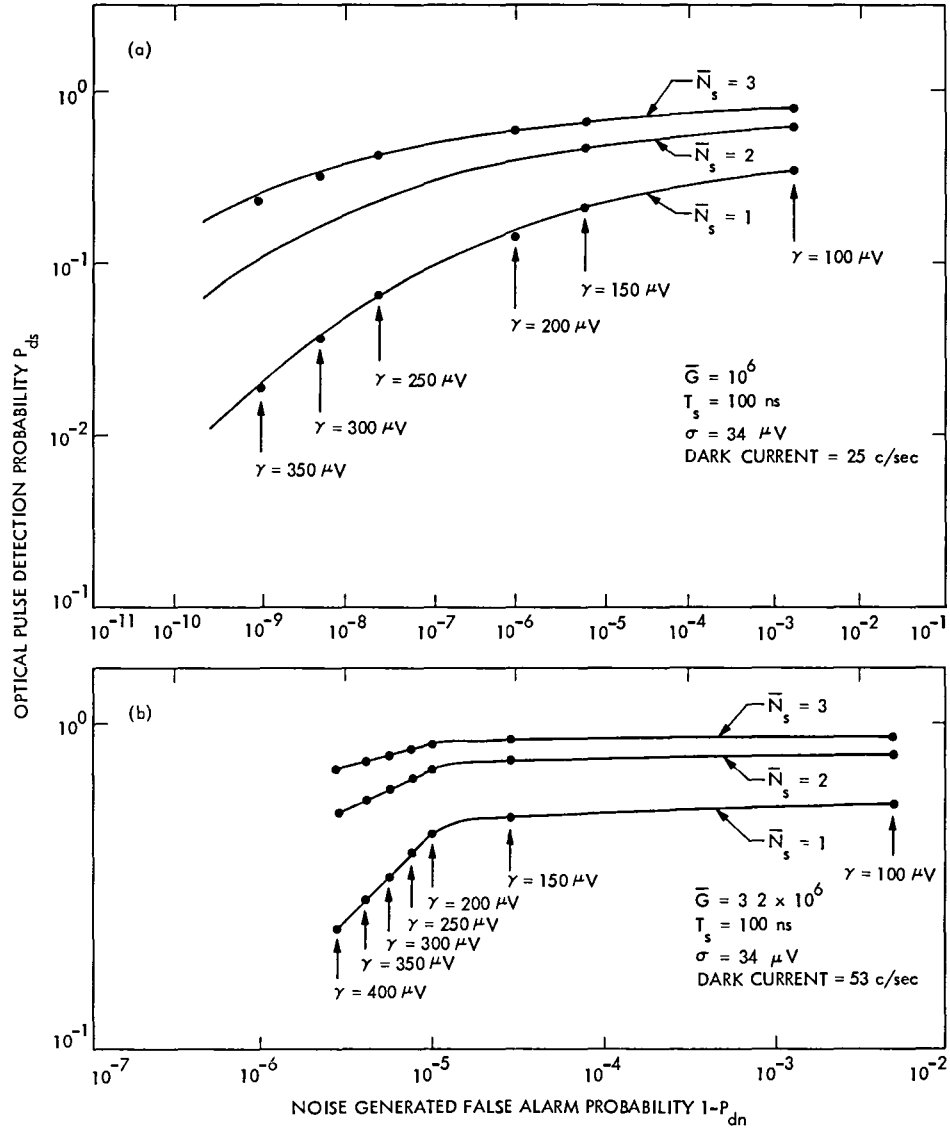


Fig. 6. Optical pulse detection and false alarm probabilities for \bar{N}_s photoelectrons emitted per pulse: (a) PMT gain = 10^6 , 25 dark counts/sec, (b) PMT gain = 3.2×10^6 , 53 dark counts/sec

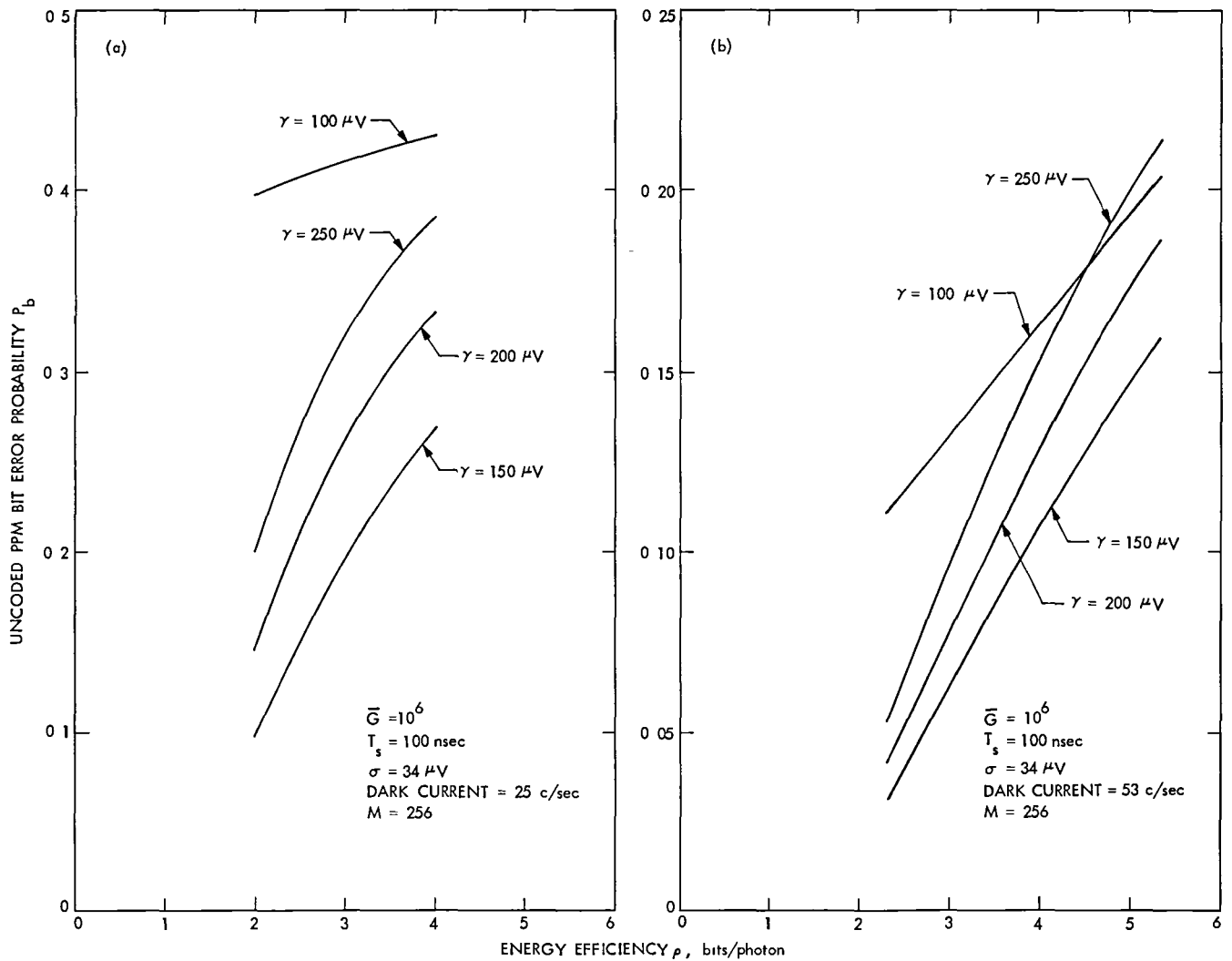


Fig. 7. Bit error performance for uncoded 256-ary PPM (a) PMT gain = 10^6 , 25 dark counts/sec, (b) PMT gain = 3.2×10^6 , 53 dark counts/sec

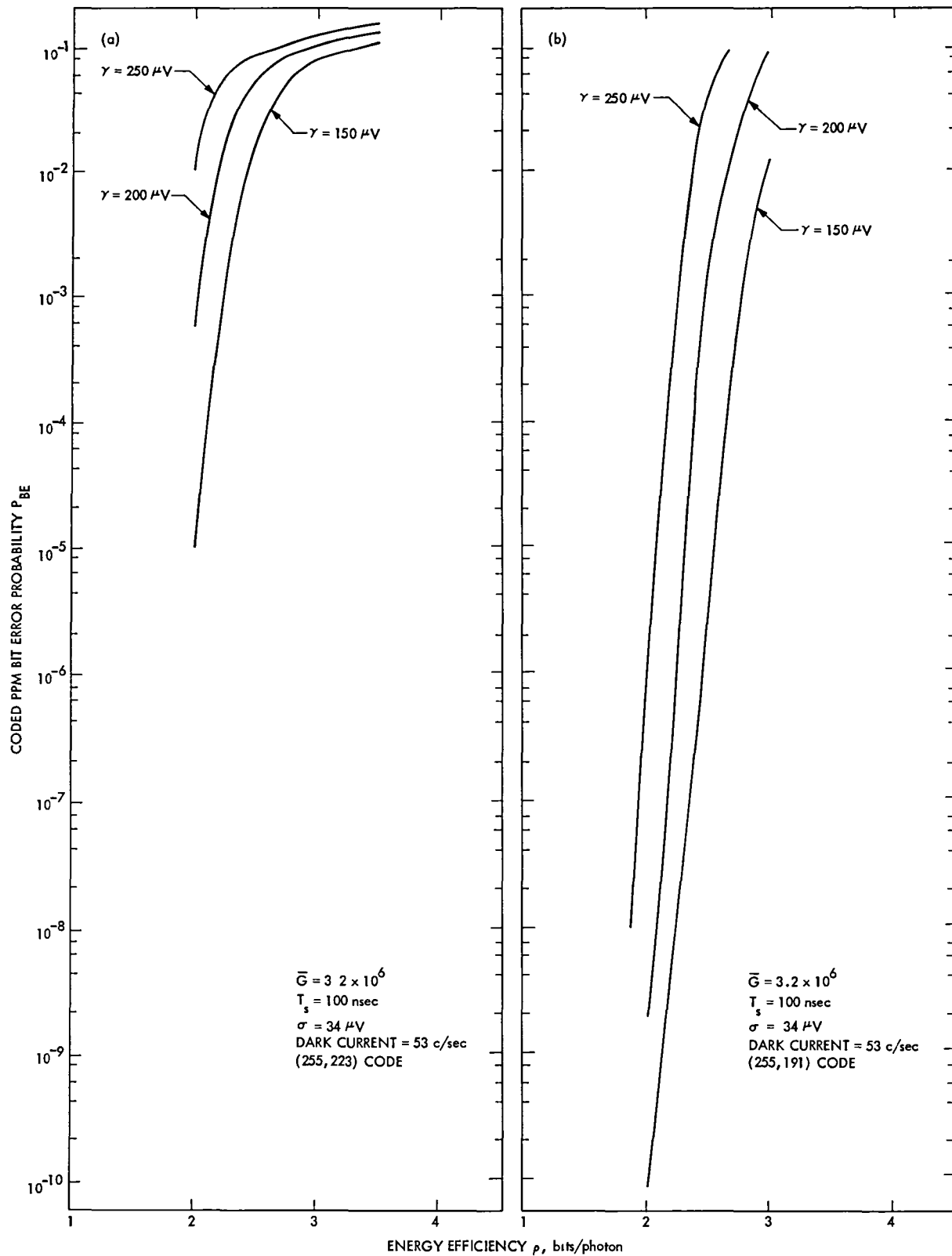


Fig. 8. Coded bit error performance for PMT gain of 3.2×10^6 and 53 dark counts/second: (a) (255,223) rate 7/8 R-S code, (b) (255,191) rate 3/4 R-S code

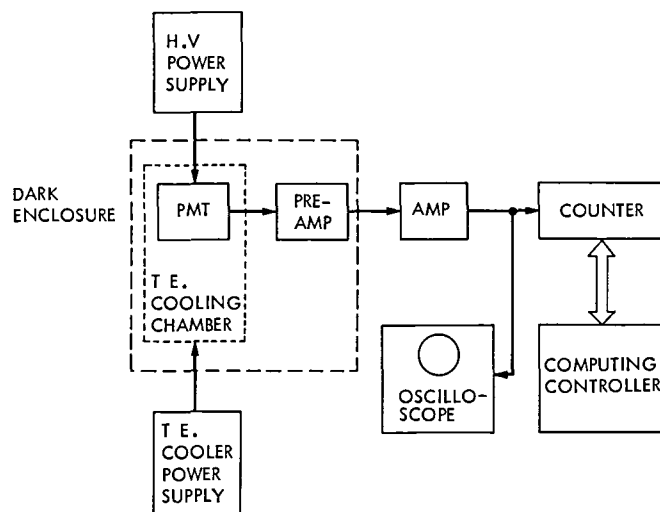


Fig. 9 Experimental setup for phase I test

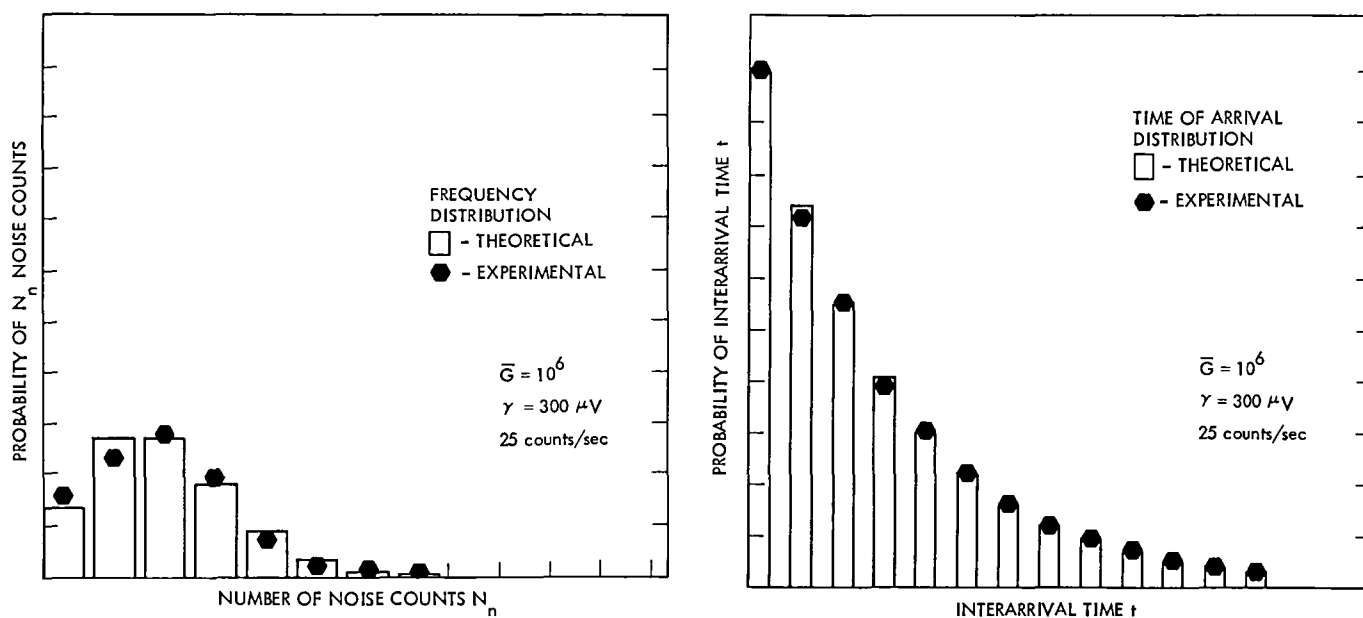


Fig. 10 Theoretical and experimental dark count noise distributions (gain = 10^6 , 25 dark counts/sec) (a) frequency distribution (theory = Poisson), (b) interarrival time distribution (theory = exponential)

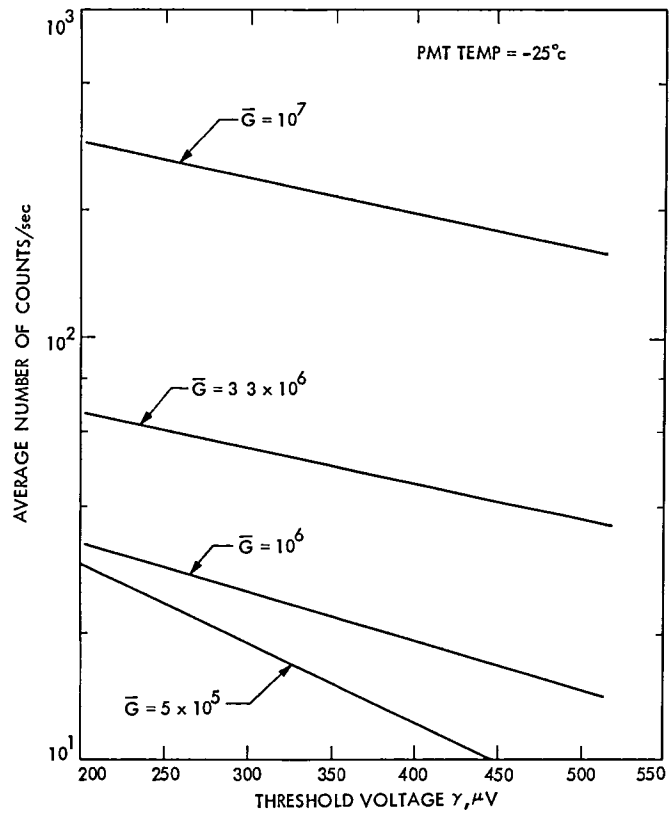


Fig 11. Measured PMT - amplifier dark current intensity

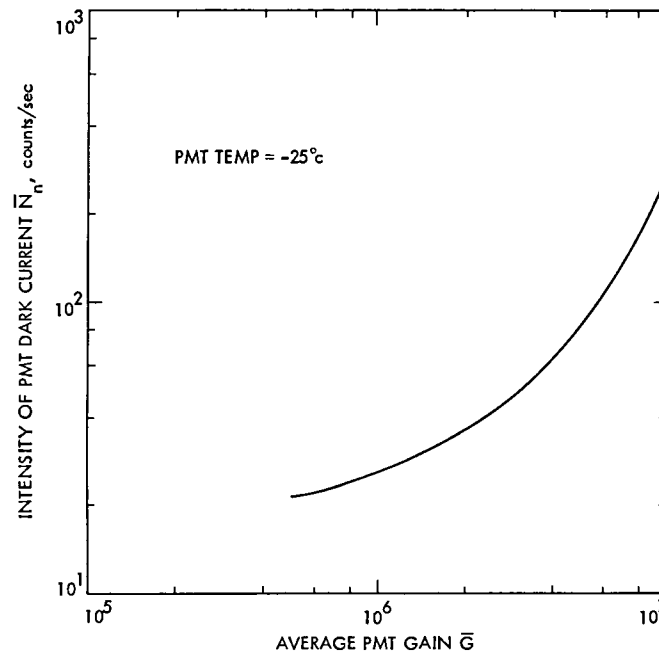


Fig. 12. Variation in PMT dark current intensity with average PMT gain

Appendix

In Section III we found that the uncoded PPM symbol error probability was given by (Eq. 17)

$$P_S = 1 - P_{ds}(P_{dn})^{M-1} - \frac{1}{M} (1 - P_{ds}) P_{dn}^{M-1} - \sum_{k=1}^{M-1} \frac{1}{k+1} \binom{M-1}{k} P_{dn}^{M-1-k} (1 - P_{dn})^k P_{ds} \quad (\text{A-1})$$

Combining the second and fourth terms and changing the index of summation yields

$$P_S = 1 - \frac{1}{M} (1 - P_{ds}) P_{dn}^{M-1} - \sum_{k=1}^M \frac{1}{k} \binom{M-1}{k-1} P_{dn}^{M-k} (1 - P_{dn})^{k-1} P_{ds}$$

Then multiplying and dividing the last term by $M(1 - P_{dn})$ produces

$$\begin{aligned} P_S &= 1 - \frac{1}{M} (1 - P_{ds}) P_{dn}^{M-1} \\ &\quad - \frac{P_{ds}}{M(1 - P_{dn})} \sum_{k=1}^M \binom{M}{k} P_{dn}^{M-k} (1 - P_{dn})^k \\ &= 1 - \frac{1}{M} (1 - P_{ds}) P_{dn}^{M-1} - \frac{P_{ds} (1 - P_{dn}^M)}{M(1 - P_{dn})} \quad (\text{A-2}) \end{aligned}$$

which is the desired result

That this expression is correct can be seen from the following analysis.² The second term is clearly the probability of making a correct symbol decision when a PPM symbol erasure occurs (cf Eq. 13). The last term of equation (A-2) must therefore be the probability of correct decision when one or more pulses occur in the PPM symbol. Let us hypothesize a different demodulator decision rule. We will assume that erasures are still resolved by the roll of an M -sided die, but that in the event of one or more pulses in a symbol time, the first pulse is *always* taken as correct. This demodulation rule is really equivalent to the maximum-likelihood decision rule when the inputs to the demodulator are equally likely and one is interested in average probability of error (Note the actual demodulator uses a modified form of this rule.) For this new demodulator the probability of correct symbol decision given that the actual signal slot was the first slot is P_{ds} . The probability of correct decision given the second slot was actually sent is likewise $P_{dn} P_{ds}$. Continuing on in this manner and using the fact that each of the conditions has equal chance of occurring results in

$$P_r \text{ (correct decision and one or more pulse received)}$$

$$= \frac{1}{M} (P_{ds} + P_{dn} P_{ds} + P_{dn}^2 P_{ds} + \dots + P_{dn}^{M-1} P_{ds})$$

This clearly equals the last term of Eq. (A-2)

²The authors are indebted to R. Stokey for this interpretation

DSS 13 Unattended Operations Station Controller Status Report

H Cooper

Radio Frequency and Microwave Subsystems Section

In this report, a brief history of the unattended operations program is presented, followed by the status of the Phase II Station Controller implementation. The DSS 13 Station Controller is being developed as a part of the S-X Systems Development work unit under RTOP-68, Station Monitor and Control Technology. This work unit provides the system engineering for RTOP-68.

I. Brief History of the Unattended Operations Program

The first major attempt to automate a Deep Space Station was in 1975 at DSS 14. Three PDP-11/20 Minicomputers and one 8080 based microprocessor computer were used to control the Transmitter, Receiver, Microwave, and SDA subsystems. Communication between the station controller computer, the subsystem computers and the equipment itself was via the JPL standard interface. In some cases, the equipment already had the JPL interface installed, in others the interfaces had to be designed and installed.

The programs for the PDP-11 minicomputers were written in Basic, and the one for the microprocessor computer in PL/M. These programs were debugged as much as possible at JPL. The computers and interface equipment were then shipped to DSS 14 for testing.

Testing was carried out on a noninterference basis at DSS 14 over a period of about a month. Although each subsystem

performed nearly as planned, the poor reliability of the hardware and software prevented the entire system from operating successfully for any length of time.

Although this first exercise was not completely successful, it did yield valuable information applicable to the next phase of automation which was to be at DSS 13. In general, the main problem areas were found to be

- (1) Hardware reliability
- (2) Software reliability and speed
- (3) Communication and protocol
- (4) Operator interface
- (5) Project personnel control

After the DSS 14 testing was completed, it was decided to move further automation development to DSS 13. Since DSS 13 was an R&D station, the scheduling of station time for

testing, and equipment development was much easier than at DSS 14 where the station was largely committed to spacecraft tracking. It was further decided to change from using the PDP-11 minicomputers with their magtape operating systems to ROM-based microprocessor computers. These microcomputers were similar to the one used at DSS 14 which showed greater reliability than the PDP-11s.

The decision to use a microprocessor computer prompted the development of JPL's first "standard microcomputer". This microcomputer used commercially available 1st-generation modules and components. At this early stage in microprocessor development, the software, written in PL/M, had to be cross-compiled on the GPCF 1108 computer and loaded into the microprocessor computers via paper tape, or prom modules. Software development and debugging on these early systems was laborious and slow.

The initial DSS 13 system was similar to the one at DSS 14. The Receiver and Transmitter subsystems at DSS 13 had to have remote control and monitoring interfaces fabricated. Finally the Receiver, Transmitter, Microwave, and SDA systems were automated as per plan. However, as testing progressed, it became apparent that the original plan had operational deficiencies which were gradually corrected.

Prior to the moving of the unattended operation effort to DSS 13, the antenna system at DSS 13 had already been placed under computer (Modcomp minicomputer) control. In order to provide a complete tracking facility, after the RF subsystem had been completed, antenna control was added to the automation effort. Since this was an afterthought, the interface, although workable, was not a clean one.

Another major addition to the system was the capability to download prefix and configuration data, as well as remotely operating the station from JPL. This was accomplished by connecting a remote terminal, with cassette tapes, to the station controller using a high speed data line and modems.

After these improvements were added and debugged, the station was operated in an unattended mode. Delivery of data to projects for a period of six months in 1979 was accomplished while the DSS 12 antenna was being enlarged from 26 to 34 meters.

Shortly afterwards, it started becoming apparent that an upgrade of both hardware and software was desired. Since moving to DSS 13, large advances had taken place in improved microprocessor hardware and software development tools. Station operation generated new requirements in the areas of operator interface, system status reporting, and generation and transferring of prefix and configuration data. For these

reasons, it was decided to implement what is now known as the Phase II system.

II. Phase II Plan

Phase II provides the opportunity to upgrade both the hardware and software of the subsystem controllers as well as the station controller. In general, all of the subsystem controllers will be converted to a MULTIBUS chassis thereby providing the capability to select from the large number of commercially available MULTIBUS compatible modules. Although there are a large number of modules to select from, the unattended operations effort is attempting to standardize on a minimum set of modules in order to reduce spares, and to develop a thorough understanding of the operation of the common modules used. The initial module selection consists of the following:

- (1) BLC-80/204 Single board computer
- (2) BLC-064 64K RAM module
- (3) BLC-8432 32K EPROM module
- (4) BLC-8201 Floppy Disk Controller
- (5) SBC-534 4 CH RS-232 module
- (6) SBC-519 Parallel I/O module
- (7) JPL-QSIA JPL designed interface module

One of the major advantages of standardizing on the MULTIBUS chassis and modules, is that with the addition of a set of floppy disk drives, the software can be entered, compiled, debugged, and executed on the target computer in RAM memory. This eliminates the need for separate development systems and means of downloading the compiled programs. Once it has been determined that the program is operating properly in RAM memory, it can then be "burned" into PROMs for greater reliability.

Software improvements will result in improved system monitor, control, and status reporting. The communication protocol between controllers will be changed to readable ASCII and be more generalized.

III. Phase II Station Controller Hardware

The functions of the Phase II station controller (see Fig. 1) are as follows:

- (1) Facilitate remote control of the station
- (2) Provide an interface to the Information Processing Center (IPC). Through this interface, prefix and

configuration data can be transferred from the IPC to the station without operator type-ins, and pass log data can be transferred to the IPC for dissemination and distribution to users

- (3) Coordinate the operation of the various subsystems through their respective controllers

In contrast to the Phase I station controller which contained one microcomputer with a terminal at DSS 13 and a remote terminal at NOCC, the Phase II station controller actually is composed of two microcomputers linked by a 7200 baud data line. The driving factor for having a microcomputer at NOCC is to provide for improved operator interface, and an interface to the IPC.

One significant addition to the Phase II station controller is the floppy disk drives. These are now required for four reasons:

- (1) Pass log data must be stored temporarily prior to being transferred to the IPC
- (2) Configuration and prefix data from the IPC must be temporarily stored prior to transfer to the DSS-13 controller
- (3) Menus and help prompt frames, which are anticipated to be many, are too large to be stored permanently in the computer memory
- (4) A variety of support programs may need to be loaded and executed (i.e., the program which transfers files to and from the IPC is a separate program from the one which controls DSS 13)

Each controller will have three CRT displays. One of these will be reserved for real time system status, another CRT will be reserved for the display of routine log messages from the subsystem controllers, and the third CRT will be reserved for operator interaction along with the keyboard. The printer is an optional utility which may be used for listing log messages or any other desired file.

Each of the controllers has enough peripherals to function as a software development system. All software entry, compilation, debugging and execution can be performed on either of the station controller computers.

IV. Phase II Microcomputer

Both of the microcomputers used in the station controller are identical with the exception of the JPL-QSIA (Quad Standard Interface Adapter) module located in the DSS 13 controller (see Fig 2). There is no need for this module in

the NOCC controller as the NOCC controller does not interface with any hardware having the JPL standard interface.

The microcomputer chassis is a standard MULTIBUS chassis containing commercially available modules with the exception of the specially designed JPL-QSIA module. This computer is configured as a multiprocessor computer in that there are six modules which have microprocessor devices on them. The function of the modules are described below:

1 BLC-80/204 These modules are the basic single board computers which run the application programs. One of these modules is dedicated to performing the protocol requirements of the high speed data line between the two controllers and its program is permanently "burned" into PROMs. The other single board computer performs the primary function of the microcomputer.

2 BLC-8201 This module is a floppy disk controller. It controls two disks in single density mode, each disk having a capacity of approximately 250K bytes.

3 BLC-8229 These modules are CRT/KEYBOARD controllers. One is required for each of the CRTs. The keyboard is connected to only one of the modules. Although these modules have the capability of executing general subroutines, it is anticipated that they will not be used for that purpose as it would complicate the program unnecessarily.

4 SBC-534 This module is a four channel RS-232 module. It supports a variety of baud rates and is used to interface the printer, the IPC modem, and to receive serial station time.

5 BLC-064 This is a 64K RAM memory module.

6 JPL-QSIA This module is a four channel interface designed to operate with any equipment equipped with a JPL standard interface, such as the star switch to which it is connected. Other subsystem microcontrollers use this module to interface to the SDAs, the Receivers, and etc.

Since this microcomputer has six "master" modules (BLC-80/204s, BLC-8229s, and the BLC-8201), the normal serial priority bus arbitration scheme must be replaced by a parallel bus arbiter. This function consisting of two IC's is wired on a prototype wirewrap module.

The NOCC Phase II hardware is configured in a five bay wraparound console with a desk top. The displays and keyboard are positioned for best operator viewing and use. The DSS 13 Phase II hardware is presently installed in a standard DSN rack but will be installed in a similar five bay console shortly. The DSS 13 console will be shared with the antenna.

controller and display. All of the Phase II station controller hardware has been installed and tested.

V. Phase II Station Controller Software

The present software residing in the Phase II controller is a slight modification of the previous Phase I software, primarily because the subsystem controllers have not all been converted to the Phase II hardware configuration, and do not have the Phase II communication and message protocol installed.

The program residing in the NOCC controller for automatically dialing and accessing data files in the IPC has been written and tested, similarly, the transfer of pass logs to the IPC can also be accomplished. Programs which will reside in

the IPC that generate the predix and configuration files, and disseminate the pass log data that has not been written.

VI. Future Plans

Future plans for the station controller include the following in order of importance:

- (1) Complete the Phase II software sufficient to perform an uplink demonstration in late March 1982.
- (2) Complete the Phase II software.
- (3) Integrate touch screen, voice recognition, and voice response systems for improved operator control.
- (4) Complete the software residing in the IPC for predix generation and pass log dissemination.

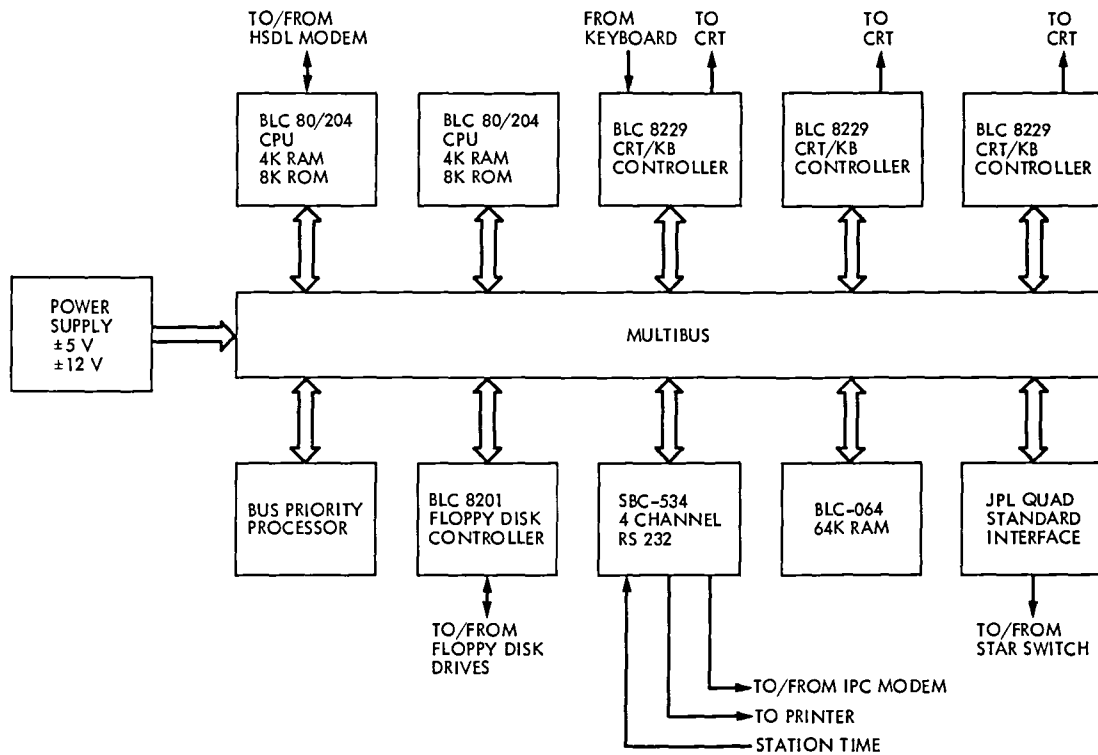


Fig 1. NOCC/DSS 13 microcomputer block diagram

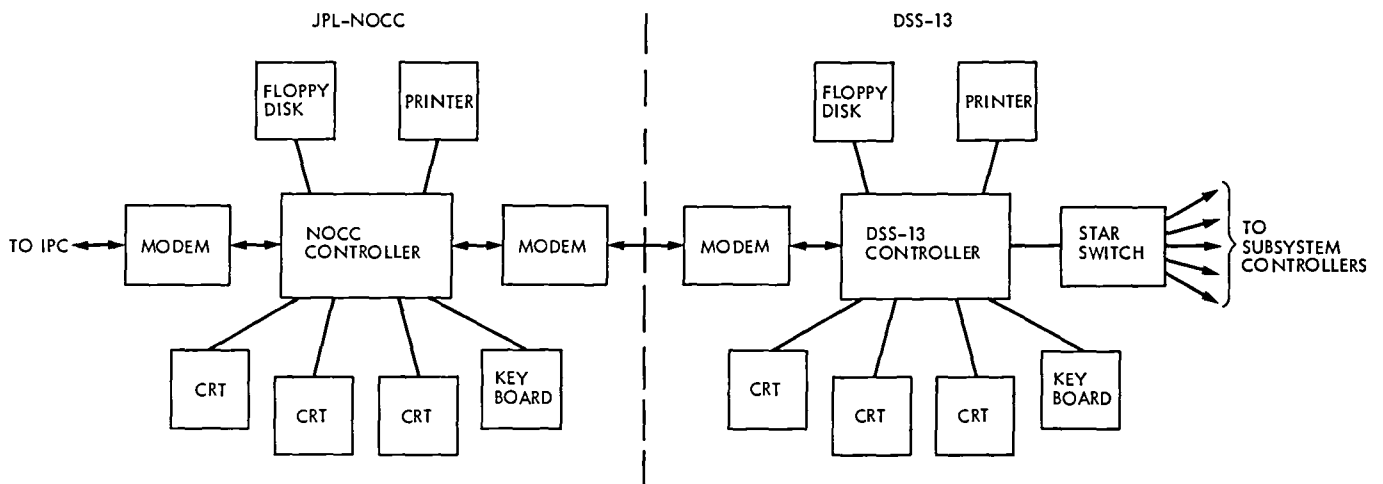


Fig 2. Phase II station controller

A Digital Filter Implementation of the Deep Space Transponder Phase Lock Loop Integrator

J W Heller

Communications Systems Research Section

The phase lock loop integrating filter in the NASA standard deep space transponder has suffered in reliability due to the malfunctioning of feedback capacitors. In this paper an alternative design, using a digital filter, is presented which eliminates these critical components. This filter can also be used in other flight and ground equipment utilizing phase lock loops.

I. Introduction

In the present version of the NASA deep space transponder the phase lock loop integrating filter has an extremely long time constant of 3000 seconds. The electronic circuit which realizes this consists of an operational amplifier with a large (11.6 μ f) feedback capacitance. These capacitors are critical as any leakage current can drastically affect the pole location of the desired transfer function. Unfortunately, the capacitors have not been extremely reliable, causing a degradation of the loop performance. This article discusses a digital implementation of the integrating filter and, although capacitors are not eliminated, the capacitances are much smaller and used in an antialiasing filter where the filter requirements are less critical.

II. Analysis of the Linear Integrating Filter

The large feedback capacitance is a consequence of the specified filter transfer function

$$F(s) = \frac{1 + \tau_2 s}{1 + \tau_1 s} \quad (1)$$

where

$$\tau_2 \approx \frac{4d^2 + 1}{2W}$$

and

$$\tau_1 \approx \frac{\tau_2^2}{4d^2} \cdot A \cdot K$$

where d is the loop damping, W is the loop noise bandwidth, K is the open loop gain and A is the rms signal amplitude. For the deep space transponder $\tau_2 \approx 0.083$ seconds and $\tau_1 \approx 3000$ seconds

A differential input integrating filter circuit similar to the one in the transponder is shown in Fig 1. The transfer function for this circuit is

$$e_o = - \frac{\frac{R_F}{R_2} \left(\frac{1}{sC} + R_1 \right)}{1 + \frac{1}{A} + \frac{R_F}{AR_2} \left(\frac{1}{sC} + R_1 + R_F \right)} \cdot e_i \quad (2)$$

Since A is large ($3 \cdot 10^5$) and the ratio R_F/AR_2 is small, then Eq (2) simplifies to

$$e_o \approx - \frac{R_F}{R_2} \frac{\left(\frac{1}{sC} + R_1 \right)}{\left(\frac{1}{sC} + R_1 + R_F \right)} e_i \quad (3a)$$

This equation can be rearranged to resemble Eq (1)

$$\frac{e_o}{e_i} = - \frac{R_F}{R_2} \frac{1 + R_1 Cs}{1 + (R_1 + R_F) Cs} \quad (3b)$$

and by comparison with Eq (1)

$$\tau_2 = R_1 C \quad (4)$$

and

$$\tau_1 = (R_1 + R_F) C \quad (5)$$

For $\tau_1 = 3000$ seconds Eq (5) implies a large value for C . For example, in the present transponder filter design R_F is realized using a T-network (not shown in Fig 1) to achieve a

value of $260 \text{ M}\Omega$. However, even this resistance summed with the $7.5 \text{ k}\Omega$ resistance chosen for R_1 requires a steep capacitance of $11.6 \text{ }\mu\text{f}$ which, in practice, requires two $5.8\text{-}\mu\text{f}$ capacitors in parallel. Unfortunately, even small leakage currents in these capacitors can produce large changes in the circuit transfer function. Therefore alternative designs which reduce or eliminate the capacitors are worth considering. One such design using a digital filter performs the integration using binary additions and multiplications. No large capacitances are required, and where capacitors are used their leakage effect on the transfer characteristic is much less than for the present circuit.

III. Digital Filter Design

In order to preserve the transient response of the linear integrator an impulse-invariant, infinite impulse response, digital filter structure was derived as follows. The impulse response of the linear filter

$$h(t) = \mathcal{L}^{-1} [H(s)] = \mathcal{L}^{-1} \left[\frac{1 + \tau_2 s}{1 + \tau_1 s} \right]$$

$$= \frac{\tau_2}{\tau_1} \delta(t) + \left(\frac{1}{\tau_1} - \frac{\tau_2}{\tau_1^2} \right) e^{-t/\tau_1} u(t) \quad (6)$$

was set equal to the desired digital filter output at each sampling interval, i.e.

$$h(n) = h(t) \Big|_{t \rightarrow nT} \quad (7)$$

Then the z transform of the sampled impulse response was determined

$$H(z) = Z[h(n)] = \frac{\left(\frac{\tau_2 + 1}{\tau_1} - \frac{\tau_2}{\tau_1^2} \right) - \frac{\tau_2}{\tau_1} e^{-T/\tau_1} z^{-1}}{1 - z^{-1} e^{-T/\tau_1}}$$

and from this system function the digital filter structure can be drawn (Fig 2). The choice of the sampling interval T depends on the precision with which the coefficients can be represented in the filter hardware (i.e., register width) and the antialiasing filter requirements. The latter is needed to attenuate signal frequencies above the Nyquist rate, $f_s/2 (=1/2T)$.

IV. Implementation

Figure 3 shows the schematic diagram for the digital filter and its associated circuits. The left-most operational amplifier is used for scaling and retains the differential input configuration of the original linear filter. The second op-amp serves as an antialiasing filter using a 3-pole Butterworth low-pass filter design. The cutoff frequency of 1 Hz was chosen so that aliasing from frequencies greater than 10 Hz are attenuated by at least 54 dB and therefore are less than the least significant bit of the digital filter's 9 bit digital-to-analog converter.

The analog-to-digital conversions, digital filtering and digital-to-analog conversion are all performed by an Intel 2920 digital signal processing integrated circuit. A functional diagram is shown in Fig. 4. The 2920 is the most compact of the digital signal processors available as it incorporates a digital-to-analog converter (but not the voltage reference) on the chip along with the usual digital components such as an ALU, registers, etc. The filter structure is realized in software stored in an erasable memory capable of holding 192 instructions. Approximately 35 instructions are needed to convert the analog signal into a 9 bit binary representation. The reverse process requires another 15 memory locations. In addition, a "correction curve" was programmed to minimize the errors resulting from the nonlinear output of the 2920's sample-and-hold buffer (for a discussion of this problem see Ref. 1). Also, a software timer was included in order to execute the filter code at the appropriate submultiple frequency of the 2920 hardware clock. The submultiple frequency was set

equal to the sampling frequency of 20 Hz. Figure 5 shows an excerpt of the 2920 code listing the digital filter structure. Although the internal register width of the 2920 is 25 bits, this was not of sufficient accuracy for representing the pole coefficient, so much of the filter computation was done using two registers for each variable (extended precision).

The op-amp to the right of the 2920 is used as an adjustable-gain voltage follower. A fourth op-amp provides a precise reference voltage for the 2920's digital-to-analog converter.

V. Performance

The circuit was tested for its step response and frequency response. The output for a 0.75-volt step input was recorded over a 3-hour period at 5-minute intervals and compared with the ideal response to the linear integrator. The maximum error (excluding near-zero values seen during the first minute and attributable to offsets, etc.) was 1.5 percent. Frequency responses of the digital and ideal filters are shown in Fig. 6. It was not possible to test the filter in a loop due to budget restrictions.

VI. Conclusion

It appears possible to replace the present linear integrator in the NASA deep space transponder with a digital filter that is not as susceptible to capacitor reliability problems as are the filters presently used.

Reference

1. Heller, J., "An Evaluation of the Intel 2920 Digital Signal Processing Integrated Circuit," *TDA Progress Report 42-63*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 108-113, June 15, 1981.

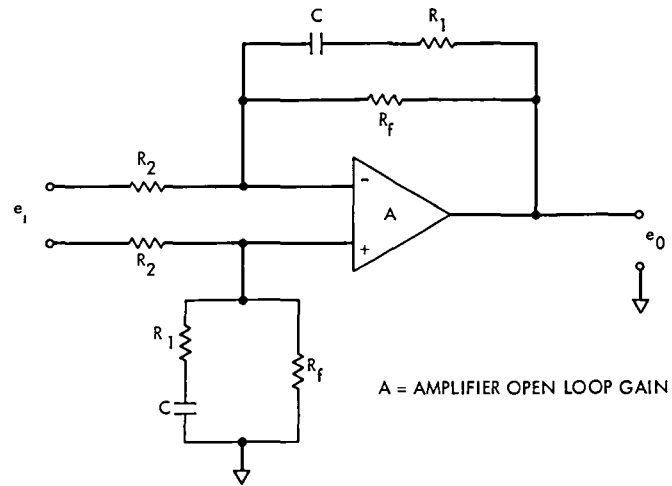


Fig 1. Simplified schematic of the deep space transponder integrator

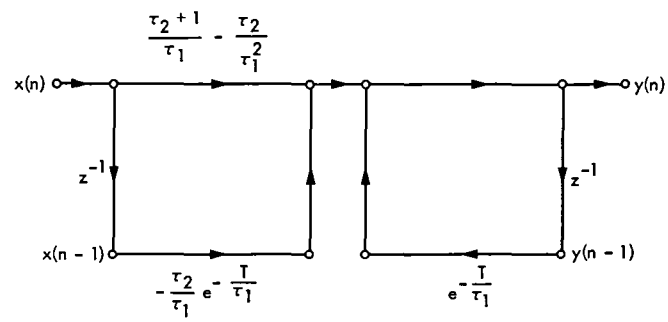


Fig. 2 Digital filter structure for Integrator

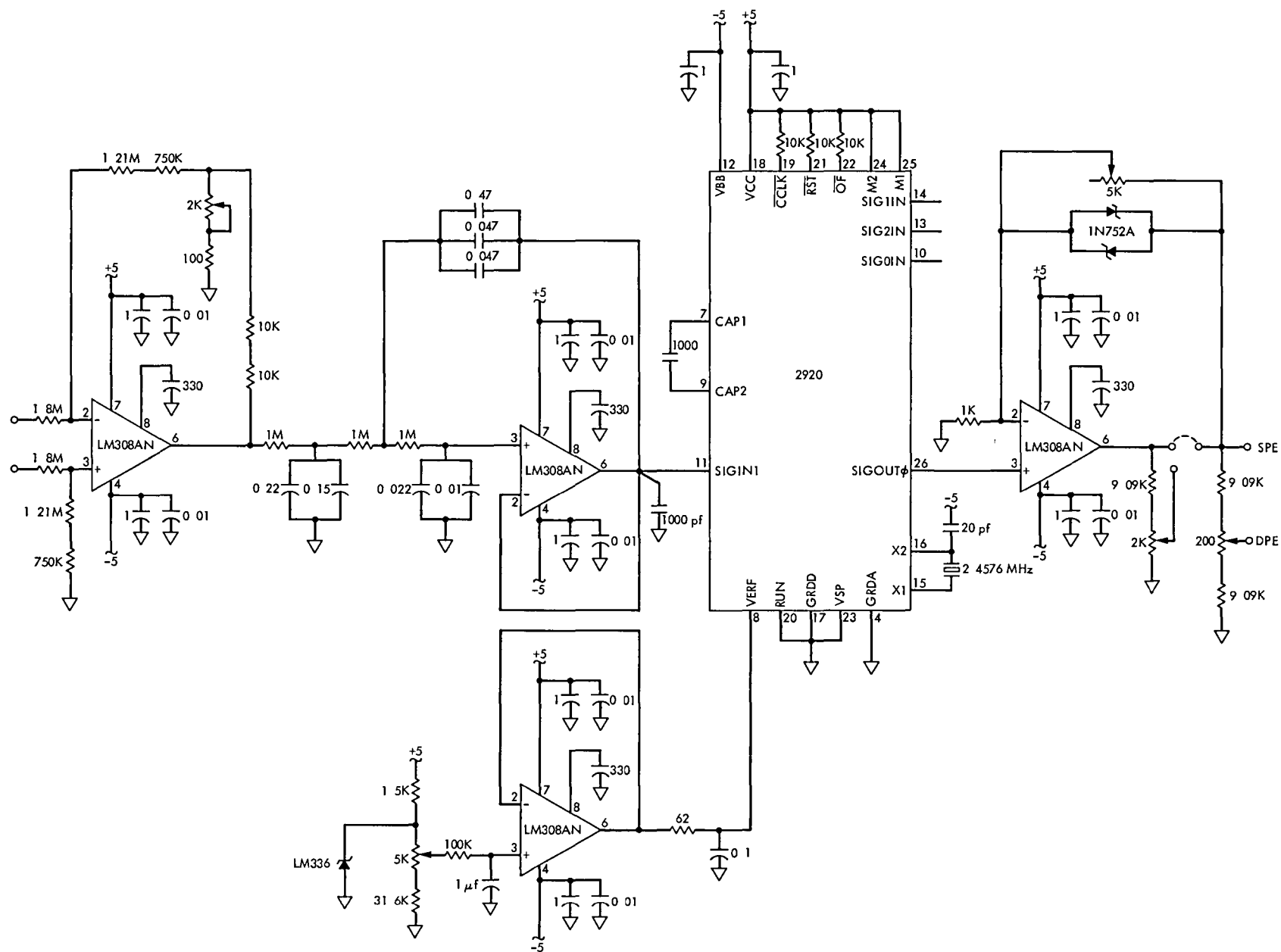


Fig 3. Digital filter schematic diagram

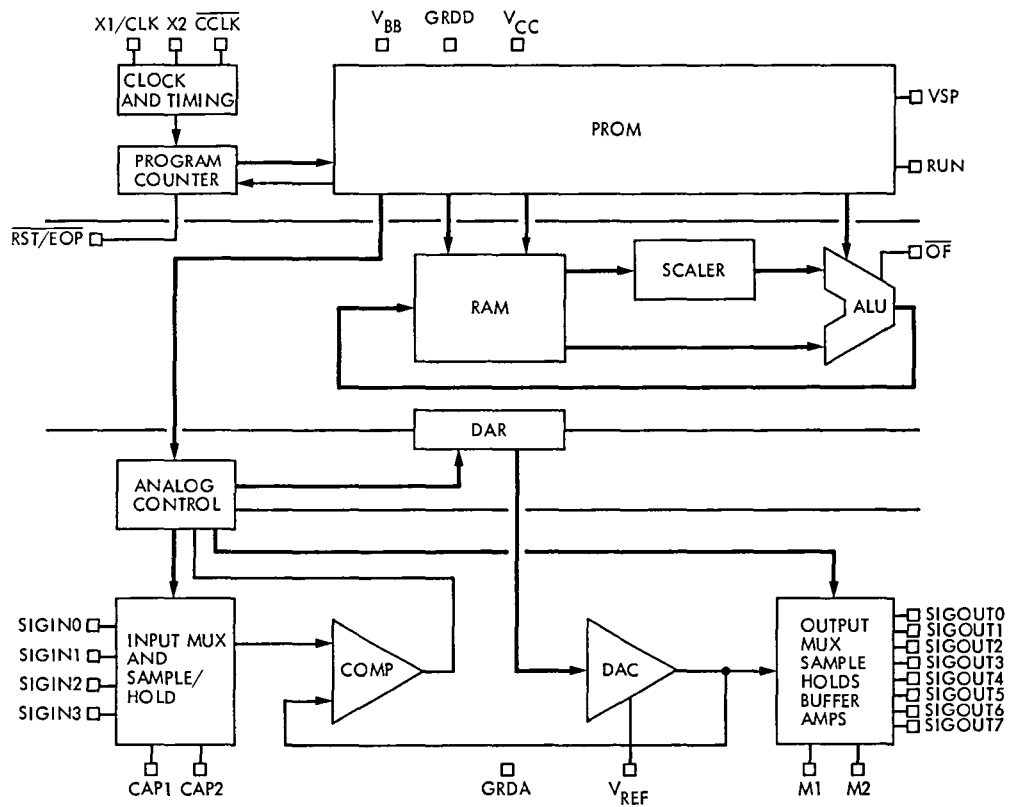


Fig 4 Block diagram of 2920 signal processor

```

LDA 0,DAP,R03,INPUT SCALED HERE
    0=0 125000000+DAP
ADD 0,0,R13
    0=0 125015258+DAP
SUB 0,DAP,R03
    0=1 52500071 10+15+DAP
ADD 0,0,R02
    0=1 9071509 10+15+DAP
ADD 0,0,R06
    0=1 9371533 10+15+DAP
SUB 0,0,R04
    0=1 8160911 10+15+DAP
.END OF MACRO ADDCON
.
SUB MULTIPLE SAMPLING TIMEP
.
SUB CNT,R01,R06,SUBTRACT 0 00000001 FROM CNT
LDA DAP,CNT,R00,MOVE COUNT (CNT) TO DAP TO CHECK SIGN
LDA CNT,R06,R00,CNDS IF NEGATIVE COUNT RE-INITIALIZE TO 255
ADD CNT,R06,R00,CNDS CONTINUE BUILDING INITIALIZATION SUM
NOP
.
LOW-PASS FILTER
.
POLE SECTION
.
UPDATE INPUT AND DELAY EVERY 1/20 TH OF A SECOND

LDA OUT0_P1,OUT0_P1,R00,CNDS UPDATE DELAY (UPPER WORD)
LDA OUT1_P1,OUT0_P1,R00,CNDS UPDATE DELAY (LOWER WORD)
LDA IN0_P1,0,R00,CNDS UPDATE INPUT

LOWER MULTIPLICATION
.
LDA OUT0_P1,OUT1_P1,R06
    OUT0_P1=0 315625000+OUT1_P1
SUB OUT0_P1,OUT1_P1,R01
    OUT0_P1=-0 48437500+OUT1_P1
SUB OUT0_P1,OUT0_P1,R07
    OUT0_P1=-0 42059022+OUT1_P1
SUB OUT0_P1,OUT1_P1,R04
    OUT0_P1=-0 54209062+OUT1_P1
ADD OUT0_P1,OUT0_P1,R12
    OUT0_P1=-0 54322332+OUT1_P1
.
UPPER MULTIPLICATION
.

```

Fig 5 Excerpt from digital filter instruction sequence

```

LDA OUT0_P1,OUT1_P1,R0E
    OUT0_P1=0 3156250000+OUT1_P1
SUB OUT0_P1,OUT1_P1,R01
    OUT0_P1=-0 43437500+OUT1_P1
SUB OUT0_P1,OUT0_P1,R07
    OUT0_P1=-0 45059092+OUT1_P1
SUB OUT0_P1,OUT1_P1,R04
    OUT0_P1=-0 54309092+OUT1_P1
ADD OUT0_P1,OUT0_P1,R12
    OUT0_P1=-0 54321338+OUT1_P1
.
. ADD CARRY FROM LOWER WORD TO UPPER WORD
.

LDA 0,OUT0_P1,R01
AND 0,KP4,R00
SUB OUT0_P1,0,R00
LDA 0,0,R13
ADD OUT0_P1,0,R10

. SCALE EACH PRODUCT
.
. LOWER PRODUCT FIRST
.
LDA OUT0_P1,OUT0_P1,R07
LDA OUT0_P1,OUT0_P1,R3E

. NEXT COPY UPPER PRODUCT
.
LDA COPY,OUT0_P1,R00

. THEN SCALE EACH UPPER PRODUCT

LDA OUT0_P1,OUT0_P1,R07
LDA OUT0_P1,OUT0_P1,R0E

. CREATE MASK TO COPY BITS FROM UPPER PRODUCT
.
LDA MASK,KP7,R13
ADD MASK,KP7,R10
ADD MASK,KP7,R07
ADD MASK,KP7,R04
ADD MASK,KP7,R01
LDA MASK,MASK,R03

.
. COPY BITS FROM UPPER PRODUCT AND SHIFT TO LOWER PRODUCT POSITION
.
AND COPY,MASK,R00
LDA COPY,COPY,L02

```

Fig 5(contd)

Fig 5 (contd)

```

LDA COPY COPY LAB
LDA COPY COPY LAB
LDA COPY COPY LAB
; ADD COPY TO COUNTER DOWN LOWER WORD, ADD NEW COPY TO UPPER WORD
;
ADD OUT1,PTX COPY,500
LDA 0,OUT0,PTX,500
AND 0,KEY,500
SUB OUT0,PTX,0,500
LDA 0,0,PTX
ADD OUT0,PTX 0,PTX
;
NOW ADD DELAYED VALUE TO COUNTER-500, DELAYED VALUE
;
ADD OUT0,PTX,OUT1,PTX,500
LDA 0,OUT0,PTX,500
AND 0,KEY,500
SUB OUT0,PTX,0,500
ADD OUT0,PTX OUT1,PTX 500
LDA 0,0,PTX
SUB OUT0,PTX,0,PTX
AND 0,KEY,500
;
RND FINALLY ADD INPUT SAMPLE TO ECU
;
ADD OUT0,PTX IN0,PTX 500
LDA 1,OUT0,PTX,500
AND 0,KEY,500
SUB OUT0,PTX,0,500
LDA 0,0,PTX
ADD OUT0,PTX 0,PTX
;
IN0,ZT END OUT0,PT
INT,ZT END OUT1,PT
LDA OUT0,ZT INT,ZT,500
OUT0,ZT=0,0156250000+INT,ZT
ADD OUT0,ZT,OUT0,ZT,LAB2
OUT0,ZT=0,073125000+INT,ZT
;
ADD OUT0,ZT INT,ZT,500
OUT0,ZT=0,060028125+INT,ZT
ADD OUT0,ZT OUT0,ZT,PTX
;
OUT0,ZT=0,060117224+INT,ZT
SUB OUT0,ZT OUT0,ZT,LAB1
OUT0,ZT=-0,060117224+INT,ZT

```



```

SUB OUT0_Z1 IN1_Z1 F10
  OUT0_Z1=-0.051052755*IN1_Z1
SUB OUT0_Z1 OUT0_Z1 F10
  OUT0_Z1=-0.051014533*IN1_Z1
ADD OUT0_Z1 IN1_Z1 F10
  OUT0_Z1=-0.060029831*IN1_Z1
ADD OUT0_Z1 IN0_Z1 F00
  OUT0_Z1=-0.050018821*IN1_Z1+1.00000000*IN0_Z1

```

Fig 5 (contd)

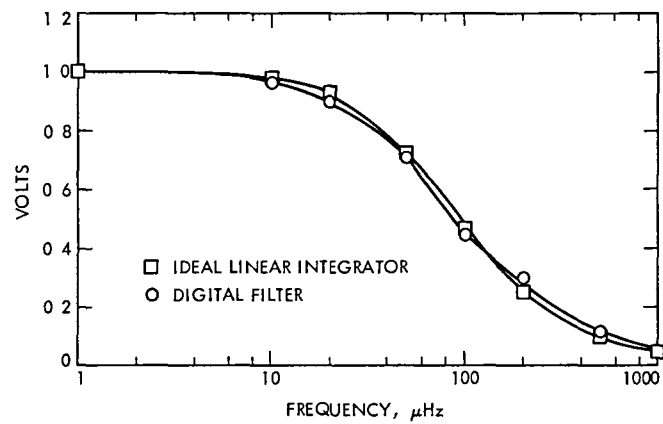


Fig. 6 Frequency response of the linear integrator and the ideal filter

Improved Carrier Tracking Performance with Coupled Phase-Locked Loops

D Divsalar

Telecommunications Systems Section

J H Yuen

Telecommunications Systems Section

and

Electrical Engineering Department,
California Institute of Technology

A carrier arraying system to combine the received carrier signals at geographically separated receivers with different carrier phases to improve carrier tracking performance is considered. This system basically couples several phase-locked loops (PLLs) to enhance received carrier signal-to-noise ratio. There is no extra alignment of carrier phases. This system automatically aligns the carrier phases, which results in coherent combining of the carrier signals. This article analyzes the tracking performance of this carrier arraying system by assessing its rms phase jitter and radio loss.

Carrier arraying alone will not provide any improvement in bit energy to noise spectral density ratio (for high carrier loop SNR). However, carrier arraying provides reduction in radio loss with respect to the no carrier arraying case. Results have been extended to the case of combined carrier and baseband arraying. Bit error rates and radio loss curves are given. Numerical results are given for three useful cases, namely, the array of 64/34-, the array of 64/34/34- and the array of 64/34/34/34-meter-antenna stations of the NASA Deep Space Network (DSN).

I. Introduction

A carrier arraying system to combine the received carrier signals at geographically separated receivers with different carrier phases to improve carrier tracking performance is considered (Ref 1). This system basically couples several phase-locked loops (PLLs) to enhance received carrier signal-to-noise ratio. There is no extra alignment of carrier phases. The system

automatically aligns the carrier phases, which results in coherent combining of the carrier signals.

This article analyzes the tracking performance of the carrier arraying system by assessing its rms phase jitter and radio loss. We have shown that the PLL in the first receiver, where the carrier arraying is performed, tracks the received carrier phase using the received carrier power from all receivers. The PLL at

station i ($i = 2, 3, \dots, N$) estimates and tracks the carrier phase difference between the received carrier phase in station i and the received carrier phase in the first station. In this system, after the carrier is acquired by the first PLL, the PLLs in the other stations track the carrier phase differences. Indeed the PLLs at the second, third, \dots , stations track the phase differences, trying to make the IF signals at the input to the carrier combiner more coherent in phase. Therefore, the loop noise bandwidth of stations 2, 3, \dots, N can be much narrower than the loop noise bandwidth of the first station. However, the first PLL should have wider loop noise bandwidth for acquisition.

Carrier arraying alone will not provide any improvement in bit energy to noise spectral density ratio (for high carrier loop SNR). However, carrier arraying provides reduction in radio loss with respect to the no carrier arraying case. Results have been extended to the case of combined carrier and baseband arraying. Bit error rates and radio loss curves are given. Numerical results are given for three useful cases, namely, the array of 64-34(T/R), the array of 64-34(T/R)-34(L/O) and the array of 64-34(T/R)-34(L/O)-34(L/O) meter-antenna stations of the NASA Deep Space Network (DSN).

II. System Model

A system for carrier and baseband arraying is shown in Fig 1. With switches SW2, SW3, \dots , SWN in their closed position, we have the combined carrier and baseband arraying system. With the switches in their open position, we have the carrier arraying only. The purpose of carrier arraying is to combine the received IF carrier signals in order to improve the carrier loop SNR. Consider the received signal from station i ($i = 1, \dots, N$), this is an RF carrier which is phase-modulated by a squarewave subcarrier ($\sum \sin \omega_{sc} t$) at a peak modulation index θ_m . The subcarrier is bi-phase-modulated with a binary data stream $D(t)$. The received RF telemetry signal from station i can be expressed as

$$S_{1i}(t) = \sqrt{2P_i} \sin [\omega_c t + \theta_i + D(t + \tau_i) \theta_m \sum \sin (\omega_{sc} t + \theta_{sci})] + n_{1i}(t) \quad (1)$$

where

ω_c is the carrier radian frequency

ω_{sc} is the subcarrier radian frequency

P_i is the total received power at station i , $i = 1, \dots, N$

θ_i is the carrier phase at station i , $i = 1, \dots, N$

θ_{sci} is the subcarrier phase at station i , $i = 1, \dots, N$

τ_i is the group delay at station i , $i = 1, \dots, N$

$n_{1i}(t)$ is the received Gaussian noise process

After coherent demodulation of $S_{1i}(t)$ by the VCO reference signal of station 1

$$r_1(t) = \sqrt{2} \cos (\omega_{LO} t + \hat{\theta}_1) \quad (2)$$

where

ω_{LO} is the VCO radian frequency of station 1

$\hat{\theta}_1$ is the phase estimate of θ_1

we get the first IF signal at station i as

$$S_{2i}(t) = \sqrt{P_i} \sin (\omega_{IF1} t + \phi_1 + \theta_i - \theta_1) \cos \theta_m + \sqrt{P_i} \cos (\omega_{IF1} t + \phi_1 + \theta_i - \theta_1) \cdot \sin \theta_m D(t + \tau_i) \sum \sin (\omega_{sc} t + \theta_{sci}) + n_{2i}(t) \quad (3)$$

where

$$\omega_{IF1} = \omega_c - \omega_{LO}$$

$$\phi_1 = \theta_1 - \hat{\theta}_1$$

Next the ambiguity due to the phase differences between stations should be resolved. This is done by phase-locked loops at stations 2, 3, \dots, N . Demodulating $S_{21}(t)$ by the reference signal

$$R_1(t) = 2 \cos (\omega_{R1} t) \quad (4)$$

and demodulating $S_{2i}(t)$ by the VCO reference signal

$$r_i(t) = 2 \cos (\omega_{R1} t + \hat{\theta}_i) \quad i = 2, 3, \dots, N \quad (5)$$

we get

$$\begin{aligned}
S_{3i}(t) = & \sqrt{P_i} \sin(\omega_{IF2} t + \phi_1 + \phi_{ei}) \cos \theta_m \\
& + \sqrt{P_i} \cos(\omega_{IF2} t + \phi_1 + \phi_{ei}) \\
& \cdot \sin \theta_m D(t + \tau_i) \sin(\omega_{sc} t + \theta_{sci}) \\
& + n_{3i}(t) \quad i = 1, 2, \dots, N
\end{aligned} \quad (6)$$

where

$$\begin{aligned}
\omega_{IF2} &= \omega_{IF1} - \omega_{R1} \\
\phi_{ei} &= \theta_i - \theta_1 - \hat{\theta}_i \quad i = 2, \dots, N \\
\phi_{e1} &= 0
\end{aligned}$$

To combine the second IF carrier signals, weight each signal by β_i , $i = 1, \dots, N$ (assume $\beta_1 = 1$) and add the signals. At the output of the IF carrier combiner at station 1, without loss of generality ignoring the component of the signal $S_{3i}(t)$ which contains data, we get

$$\begin{aligned}
S(t) = & \sum_{i=1}^N \sqrt{P_i} \beta_i \cos \theta_m \sin(\omega_{IF2} t + \phi_1 + \phi_{ei}) \\
& + \sum_{i=1}^N \beta_i n_{3i}(t)
\end{aligned} \quad (7)$$

Clearly the input signals to the carrier combiner are not in phase. But it will be shown later that when

$$\frac{B_{Li}}{B_{L1}} \ll 1 \quad (8)$$

where

B_{Li} is loop noise bandwidth of PLL at station i $i = 2, 3, \dots, N$

B_{L1} is loop noise bandwidth of PLL at station 1 without carrier arraying

Then with a good approximation, ϕ_{ei} with respect to ϕ_1 can be ignored, i.e.,

$$\phi_{ei} \cong 0 \quad (9)$$

Therefore, the input signals to the carrier combiner are approximately coherent in phase.

III. Carrier Phase Jitter Variance

The signal $S(t)$ is demodulated to a baseband signal by reference signal

$$R_2(t) = \cos(\omega_{IF2} t) \quad (10)$$

The resulting signal is

$$\begin{aligned}
\tilde{S}(t) = & \sum_{i=1}^N \sqrt{P_i} \beta_i \cos \theta_m \sin(\phi_1 + \phi_{ei}) \\
& + \sum_{i=1}^N \beta_i N[t, \phi_1 + \phi_{ei}]
\end{aligned} \quad (11)$$

which enters the loop filter of station 1 with transfer function

$$F_1(s) = \frac{1 + \tau_{21}s}{1 + \tau_{11}s} \quad (12)$$

Using linear phase-locked-loop theory, it can be shown (see Appendix A) that the equivalent reduced model of Fig. 1 for analyzing the phase jitter can be modeled as in Fig. 2. In this figure the closed loop transfer function $H_i(s)$ is given by

$$H_i(s) = \frac{\sqrt{P_i} K_i F_i(s) \cos \theta_m}{s + \sqrt{P_i} K_i F_i(s) \cos \theta_m} \quad (13)$$

where

$$F_i(s) = \frac{1 + \tau_{2i}s}{1 + \tau_{1i}s} \quad (14)$$

τ_{1i}, τ_{2i} are time constants of the loop (station i)

The loop damping parameter r_i for station i is given by

$$r_i = (\sqrt{P_i} K_i \tau_{2i}^2 / \tau_{1i}) \cos \theta_m \quad (14a)$$

With these notations the loop noise bandwidth B_{L_i} can be approximated as

$$B_{L_i} \cong \frac{r_i + 1}{4\tau_{2i}} \quad i = 1, 2, \dots, N \quad (15)$$

In Fig. 2, N_i is Gaussian noise process with one-sided noise spectral density N_{0i} , $i = 1, 2, \dots, N$.

From Fig. 2, variance of the phase noise ϕ_1 , is given by (for details see Appendix A)

$$\sigma_{\phi_1}^2 = \frac{1}{2\pi\eta} \int \left| \frac{H_1(s)}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{01}}{2P_1 \cos^2 \theta_m} + \sum_{i=2}^N \beta_i^2 \gamma_i^2 \frac{1}{2\pi\eta}$$

$$\cdot \int \left| \frac{H_1(s) [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{0i}}{2P_i \cos^2 \theta_m} \quad (16)$$

where

$$\gamma_i \triangleq \frac{\sqrt{P_i}}{\sqrt{P_1}} \quad (17)$$

Evaluation of integrals in (16) for $N > 2$ is very difficult, however, in practice, it is anticipated that the loop gains and loop filters for stations 2, 3, \dots, N will be the same, or approximately the same.

In this case

$$H_i(s) = H_2(s) \text{ for } i = 3, 4, \dots, N \quad (18)$$

Evaluating (16), using relation (18) and ignoring the $(B_{L2}/B_{L1})^n$ for $n \geq 2$ we get (see Appendix A)

$$\sigma_{\phi_1}^2 = \frac{r_1 B_{L1}}{(r_1 + 1) P_1 \cos^2 \theta_m} \cdot \left[\frac{1 + r_1 G + \xi (r_1 r_2 + r_2 G + (G - 1)^2)}{r_1 G^2 + r_2 \xi (r_1 G + G - 1)} N_{01} + \frac{1 + r_1 G + \xi (r_1 r_2 + 1)}{r_1 G^2 + r_2 \xi (r_1 G + G - 1)} \sum_{i=2}^N \beta_i^2 N_{0i} \right] \quad (19)$$

where

$$\xi \triangleq \frac{(r_1 + 1) B_{L2}}{(r_2 + 1) B_{L1}} \quad (20)$$

$$G \triangleq \sum_{i=1}^N \beta_i \gamma_i \quad (21)$$

If furthermore we consider the case when $B_{L2}/B_{L1} \ll 1$, then the expression for $\sigma_{\phi_1}^2$ can be reduced to

$$\sigma_{\phi_1}^2 = \frac{B_L \sum_{i=1}^N \beta_i^2 N_{0i}}{P_1 G^2 \cos^2 \theta_m} \quad (22)$$

where

$$B_L \triangleq \frac{1 + Gr_1}{1 + r_1} B_{L1} = \frac{1 + r}{4\tau_{21}} \quad (23)$$

and the new loop damping parameter for carrier arraying naturally is defined as

$$r = Gr_1 \quad (24)$$

Now let's consider the effect of bandpass limiter preceding the loop. Let B_{IF} represent the one-sided bandwidth of the IF filter (assuming the B_{IF} is the same for all stations) preceding the hard limiter. Then the input signal to noise ratio to the limiter of station 1 is

$$\rho_m = \frac{P_1 G^2 \cos^2 \theta_m}{\sum_{i=1}^N \beta_i^2 N_{0i} B_{IF}} \quad (25)$$

From Eq. (22) we can get the effective loop SNR ρ as

$$\rho = \frac{P_1 G^2 \cos^2 \theta_m}{\sum_{i=1}^N \beta_i^2 N_{0i} \tilde{B}_L \Gamma} \quad (26)$$

where Γ is limiter performance factor given by (Ref 2)

$$\Gamma \cong \frac{1 + \rho_m}{0.862 + \rho_m} \quad (27)$$

The loop damping parameter when a bandpass limiter precedes the loop is

$$\tilde{r} = \frac{\sqrt{8\tilde{\alpha}}_1}{\pi \sqrt{P_1} \cos \theta_m} \quad (28)$$

where $\tilde{\alpha}$ is suppression factor and is given by (Ref 2)

$$\tilde{\alpha} = \frac{\sqrt{0.7854\rho_m + 0.4768\rho_m^2}}{\sqrt{1 + 1.024\rho_m + 0.4768\rho_m^2}} \quad (29)$$

Then

$$\tilde{B}_L = \frac{1 + \tilde{r}}{4\tau_{21}} \quad (30)$$

Using Eqs (27) and (30) in (26) we notice that effective loop SNR is a monoton increasing function of ρ_m . Now we optimize the effective loop SNR with respect to weighting factors β_i , $i = 2, \dots, N$. But since ρ is a monoton increasing function of ρ_m equivalently we can optimize ρ_m with respect to β_i , $i = 2, \dots, N$. Note that ρ_m is a convex (\cap) function, thus we should have

$$\frac{\partial \rho_m}{\partial \beta_k} = 0 \quad k = 2, \dots, N \quad (31)$$

The solution of Eq (31) gives the optimum β_i as

$$\beta_i = \frac{\sqrt{P_i}}{\sqrt{P_1}} \frac{N_{01}}{N_{0i}} \quad i = 1, 2, \dots, N \quad (32)$$

Note that if Eq. (26) is optimized with respect to β_i , still we get the result given by (32)

Also note that for these values of β_i , we have

$$\frac{\partial^2 \rho_m}{\partial \beta_k^2} \leq 0 \quad (33)$$

which implies that ρ_m achieves the maximum value for β_i 's given by (32). Substituting the optimum β_i 's in (26) we get

$$\rho = \sum_{i=1}^N \frac{P_i \cos^2 \theta_m}{N_{0i} \tilde{B}_L \Gamma} \quad (34)$$

If

$$N_{0i} = N_{01} \text{ for all } i$$

then we get

$$G = \sum_{i=1}^N P_i / P_1 \quad (35)$$

and

$$\rho = \frac{GP_1 \cos^2 \theta_m}{N_{01} \tilde{B}_L \Gamma} \quad (36)$$

Now let's see how much improvement in loop SNR we get by carrier arraying. To do so, we should compare the loop SNR ρ given by (36) with the loop SNR ρ_1 for the single station number 1 (say 64-m DSN station) given by

$$\rho_1 = \frac{P_1 \cos^2 \theta_m}{N_{01} \tilde{B}_{L1} \Gamma_1} \quad (37)$$

where

$$\tilde{B}_{L1} = \frac{1 + \tilde{r}_1}{4\tau_{21}} = \frac{1 + \tilde{\alpha}_1 r_0 / \tilde{\alpha}_0}{4\tau_{21}} \quad (38)$$

$r_0 = 2$ is a damping parameter at threshold

$\tilde{\alpha}_0$ is a suppression factor at threshold

$\tilde{\alpha}_1$ is a suppression factor for single station 1

Γ_1 is a limiter performance factor for single station 1

The input SNR in the IF bandwidth at threshold is given by

$$\rho_{in,0} = \frac{2B_{L01}}{B_{IF}} \quad (39)$$

where B_{L01} is loop bandwidth at threshold, i.e., when

$$\frac{P_0}{N_0(2B_{L01})} = 1 \quad (40)$$

Then we can compute $\tilde{\alpha}_0$ using Eq (29), replacing ρ_{in} by $\rho_{in,0}$

The input SNR in the IF bandwidth for single station 1 is

$$\begin{aligned} \rho_{in,1} &= \frac{P_1 \cos^2 \theta_m}{N_0 B_{IF}} = ML \frac{2B_{L01}}{B_{IF}} \\ &= ML \cdot \rho_{in,0} \end{aligned} \quad (41)$$

where ML is carrier margin at single station 1 given by

$$ML = \frac{P_1 \cos^2 \theta_m}{N_0 (2B_{L01})} \quad (42)$$

The relation between loop SNR and the carrier margin at a single station is shown in Fig 3. Now, $\tilde{\alpha}_1$ and Γ_1 can be computed using Eq (29) and Eq (27), respectively, replacing ρ_{in} by $\rho_{in,1}$. Then the improvement in loop SNR is given by

$$\text{Improvement in loop SNR} = \rho/\rho_1 = \frac{G(\tilde{\alpha}_0 + 2\tilde{\alpha}_1)\Gamma_1}{(\tilde{\alpha}_0 + 2\tilde{\alpha})\Gamma} \quad (43)$$

Figure 4 shows the improvement in loop SNR vs carrier margin for arraying 64/34-, 64/34/34- and 64/34/34/34-m antennas, with the assumption that

$$\frac{B_{L01}}{B_{L0i}} \gg 1 \quad i \geq 2 \quad (44)$$

is satisfied. Obviously the improvement in loop SNR will be less if this assumption is not satisfied. In this case Eq (19) can be used to compute the effect of ratios of loop bandwidths. In Fig 5 we have plotted the RMS phase jitter vs carrier margin.

At this point it is interesting to note that the formula for improvement in loop SNR (43) is for loop preceded by band-pass limiter. However, if we consider second-order loop alone, the improvement will be given as

$$\text{Improvement in loop SNR} = \frac{G(1+r_1)}{1+Gr_1} \quad (45)$$

r_1 is given by Eq (14a). The formula given by (45) shows much less improvement than second order loop preceded by bandpass limiter. More interesting is if we consider the first-order loop. This case results in no improvement (zero dB improvement) with assumption (44) and even loss if assumption (44) is not satisfied. The explanation for these cases is simple. We note that by carrier arraying when assumption (44) is satisfied, the carrier power to noise spectral density will be improved by a factor of G . However, on the other hand, the loop bandwidth will be increased by a factor of

G for 1st-order loop

$\frac{1+Gr_1}{1+r_1}$ for 2nd-order loop

$\frac{1+\tilde{r}}{1+\tilde{r}_1}$ for 2nd-order loop preceded by bandpass limiter

Now let's analyze and derive the variance of the phase error ϕ_{et} . From Fig 2, it can be shown that (for details see Appendix A)

$$\begin{aligned} \sigma_{\phi_{ek}}^2 &= \frac{1}{2\pi j} \int \left| \frac{H_k(s)H_1(s)}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{01}}{2P_1 \cos^2 \theta_m} \\ &+ \sum_{\substack{i=2 \\ i \neq k}}^N \beta_i^2 \gamma_i^2 \frac{1}{2\pi j} \int \left| \frac{H_k(s)H_1(s) [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \\ &\cdot \frac{N_{0i}}{2P_i \cos^2 \theta_m} \\ &+ \frac{1}{2\pi j} \int \left| \frac{H_k(s) + H_k(s)H_1(s) \sum_{i=2, i \neq k}^N \beta_i \gamma_i [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \\ &\cdot \frac{N_{0k}}{2P_k \cos^2 \theta_m} \end{aligned} \quad (46)$$

Since we are interested in a case when the noise bandwidth of $H_i(s)$, $i = 2, 3, \dots, N$ is much narrower than the noise bandwidth of $H_1(s)$, i.e.,

$$\frac{B_{Li}}{B_{L1}} \ll 1 \quad i = 2, 3, \dots, N \quad (47)$$

then using assumption (18), it can be shown (see Appendix A) that

$$\begin{aligned} \sigma_{\phi_{ek}}^2 = & \left(\frac{G+r_2}{1+r_2} \right) \frac{N_{01}B_{L2}}{GP_1 \cos^2 \theta_m} \\ & + \frac{2r_2^2 + (1+G)r_2 + 2G}{(1+r_2)[2(1+G)r_2 + (1-G)^2]} \sum_{i=2}^N \beta_i^2 \frac{N_{0i}B_{L2}}{GP_1 \cos^2 \theta_m} \\ & + \left[1 - 2\beta_k \gamma_k \frac{(1+r_2)G + (1+2r_2)(r_2-1)}{(1+r_2)[2(1+G)r_2 + (1-G)^2]} \right] \frac{N_{0k}B_{L2}}{P_k \cos^2 \theta_m} \end{aligned} \quad (48)$$

Now if we assume

$$N_{0i} = N_0 \quad i = 1, 2, \dots, N \quad (49)$$

the optimum β_k is

$$\beta_k = \gamma_k = \frac{\sqrt{P_k}}{\sqrt{P_1}} \quad (50)$$

Using (49) and (50), the expression for $\sigma_{\phi_{ek}}^2$ can be simplified to

$$\sigma_{\phi_{ek}}^2 = \left(\frac{1+r_2+\beta_k^2}{1+r_2} \right) \frac{N_0B_{L2}}{P_k \cos^2 \theta_m} \quad (51)$$

But since $B_{Li}/B_{L1} \ll 1$, $i \neq 1$, then

$$\sigma_{\phi_{ek}}^2 \ll \sigma_{\phi_1}^2 \quad (52)$$

Also using Schwarz inequality we get

$$\frac{|\text{COV}(\phi_{ei}, \phi_1)|}{\sigma_{\phi_1}^2} < \frac{\sqrt{\sigma_{\phi_{ei}}^2 \sigma_{\phi_1}^2}}{\sigma_{\phi_1}^2} = \frac{\sigma_{\phi_{ei}}}{\sigma_{\phi_1}} \ll 1 \quad i \neq 1 \quad (53)$$

which means that ϕ_{ei} is approximately uncorrelated from ϕ_1 . Therefore, with good approximation for future analysis we can ignore ϕ_{ei} with respect to ϕ_1 .

IV. Radio Loss for Carrier Arraying

First assume the switches SW2, SW3, ..., SWN are in their open position. In order to extract the data, signal $S_{31}(t)$ is coherently demodulated to the subcarrier frequency by the reference signal

$$R_2(t) = \cos(\omega_{IF2}t) \quad (54)$$

The resulting data signal for station 1 is $(Z(t) = Y_1(t - \tau_1))$

$$\begin{aligned} Y_1(t - \tau_1) = & \sqrt{P_1} \sin \theta_m D(t) \cos(\omega_{sc}t + \theta_{sc}) \cos \phi_1 \\ & + \tilde{n}_{31}(t) \end{aligned} \quad (55)$$

where $\tilde{n}_{31}(t)$ is a white Gaussian noise process. After subcarrier tracking (assuming perfect tracking) and demodulation we have

$$X(t) = \sqrt{P_1} \sin \theta_m D(t) \cos \phi_1 + \hat{n}_{31}(t) \quad (56)$$

where $\hat{n}_{31}(t)$ is a low-pass white Gaussian noise process. The sampled signal at the output of the integrate and dump circuit is

$$x_k = \sqrt{P_1} \sin \theta_m \cos \phi_1 a_k + n_{1k} \quad (57)$$

where n_{1k} are independent white Gaussian noise samples and a_k is a k th data symbol.

At the input of the Viterbi decoder the sample x_k is 3-bit quantized. Given ϕ_1 and a_k the signal-to-noise ratio of the sample x_k is

$$SNR = \frac{(\bar{x}_k)^2}{\sigma_{x_k}^2} \quad (58)$$

where

$$\bar{x}_k = \sqrt{P_1} \sin \theta_m \cos \phi_1 a_k \quad (59)$$

$$\sigma_{x_k}^2 = \frac{N_{01}}{2T_s} \quad (60)$$

where T_s is symbol time

Therefore, the signal-to-noise ratio is

$$SNR = \frac{(\sqrt{2P_1 T_s} \sin \theta_m \cos \phi_1)^2}{N_{01}} \quad (61)$$

Note that for a rate 1/2 convolutional code the bit energy is

$$E_{b1} = (\sqrt{2P_1 T_s} \sin \theta_m)^2 \quad (62)$$

Let $f(E_b/N_0)$ represent the bit error rate for a given bit SNR E_b/N_0 . This function of $f(E_b/N_0)$ is defined

$$f(x) = \begin{cases} \exp [-(\alpha_0 + \alpha_1 x)], & x \geq \frac{\ln(2) - \alpha_0}{\alpha_1} \\ \frac{1}{2} & , x \leq \frac{\ln(2) - \alpha_0}{\alpha_1} \end{cases}$$

where $\alpha_0 = -4.4514$ and $\alpha_1 = 5.7230$ (Ref 3)

Then the conditional bit error rate is

$$P_b(\phi_1) = f\left(\frac{E_{b1}}{N_{01}} \cos^2 \phi_1\right) \quad (63)$$

Note that ϕ_1 having probability density function

$$p(\phi_1) = \frac{e^{\rho \cos \phi_1}}{2\pi I_0(\rho)} - \pi < \phi_1 < \pi \quad (64)$$

where ρ is defined by (34), then the bit error rate is

$$P_b = 2 \int_0^{\pi/2} P_b(\phi_1) p(\phi_1) d\phi_1 + 2 \int_{\pi/2}^{\pi} [1 - P_b(\phi_1)] p(\phi_1) d\phi_1 \quad (65)$$

Bit error rate performance curves for various cases are shown in Figs 6 through 8. Radio loss curves are shown in Figs 9 and 10.

In combined carrier and baseband arraying (when switches SW2, SW3, , SWN are closed), at the output of the baseband combiner we have (Ref 4)

$$Z(t) = \sum_{i=1}^N \beta_i Y_i(t - \tau_i) \quad (66)$$

Similarly, going through Eqs (56)–(62), we get

$$P_b(\phi_1) = f\left(\sum_{i=1}^N \frac{E_{bi}}{N_{0i}} \cos^2 \phi_1\right) \quad (67)$$

Then using (67) in (65), we have the bit error rate performance for the combined carrier and baseband arraying case. Bit error rate curves for the combined carrier and baseband arraying are shown in Figs 11 through 13. The Radio loss curves for this case are essentially the same as were shown in Figs. 9 and 10. Radio loss curves for combined carrier and baseband arraying, and for baseband arraying alone (Ref 4), are compared in Fig 14.

V. Conclusion

Performing carrier arraying alone reduces the radio loss of the telemetry system, but it will not improve the bit energy to noise spectral density for high loop SNR.

Baseband arraying (Ref 4) alone improves bit SNR and reduces Radio loss by a small amount, which is less than that by carrier arraying.

In order to improve bit SNR and reduce Radio loss simultaneously, the combined carrier and baseband arraying should be used.

Acknowledgement

The authors thank Marvin Simon and Milt Brockman of the Jet Propulsion Laboratory for their helpful discussions

References

1. Brockman, M , "Radio Frequency Carrier Arraying for High Rate Telemetry Reception," *DSN Progress Report 42-45*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1978, pp. 209-223.
2. Lindsey, W. C , and Simon, M K , *Telecommunication System Engineering* Prentice-Hall, Inc , Englewood Cliffs, N J , 1973.
- 3 Madsen, B , Voyager Project unpublished notes, March 1981.
- 4 Divsalar, D , Hansen, D , and Yuen, J H , "The Effect of Noisy Carrier Reference on Telemetry with Baseband Arraying," *TDA Progress Report 42-63*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1981, pp 128-135

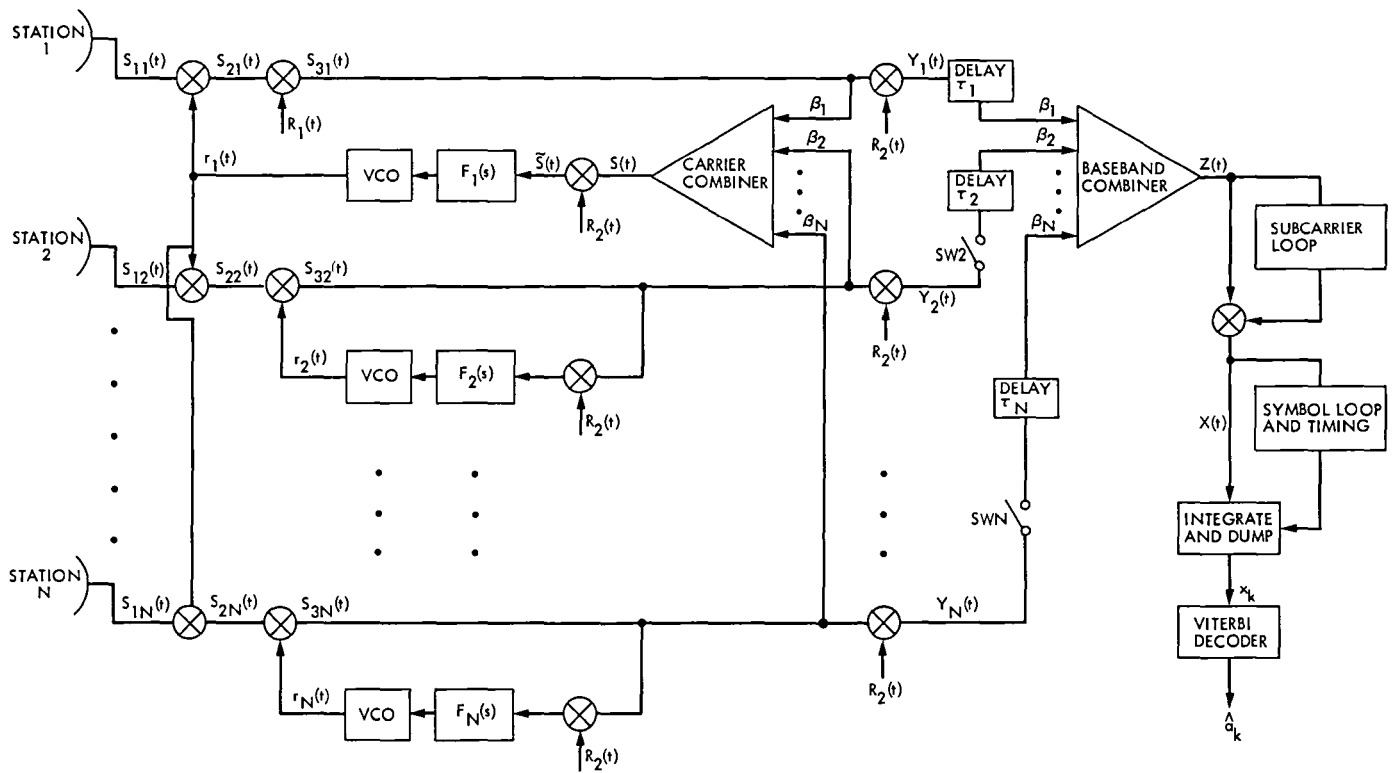


Fig 1 Configuration for arrayed network with carrier and baseband arraying

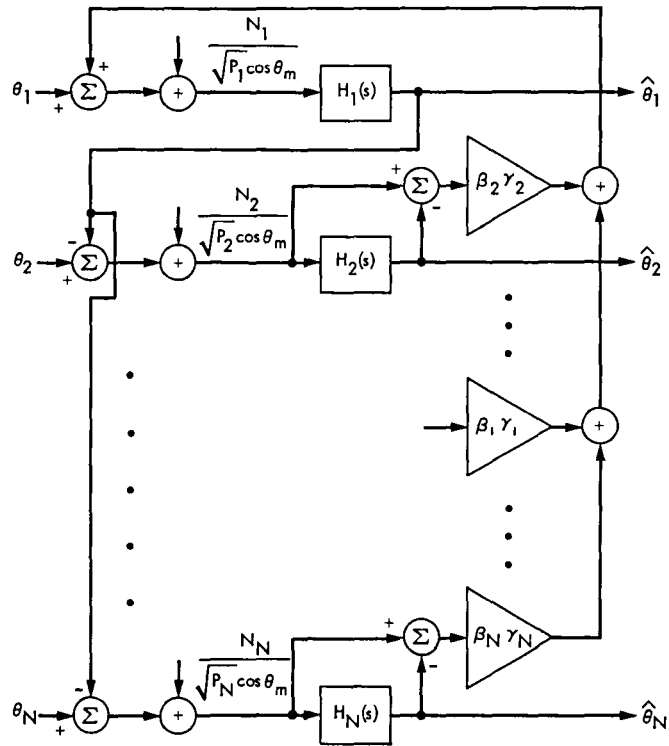


Fig. 2 An equivalent linear representation of carrier arraying system for phase jitter analysis

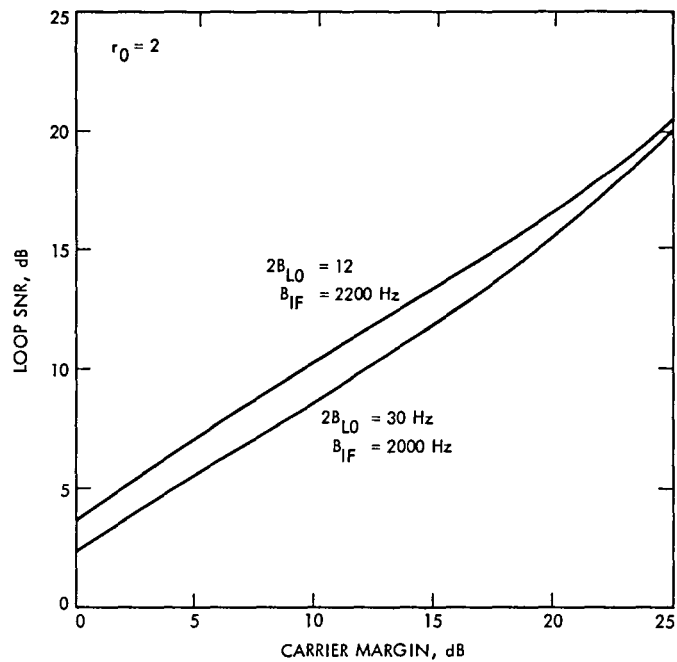


Fig. 3. Relation between loop SNR and carrier margin

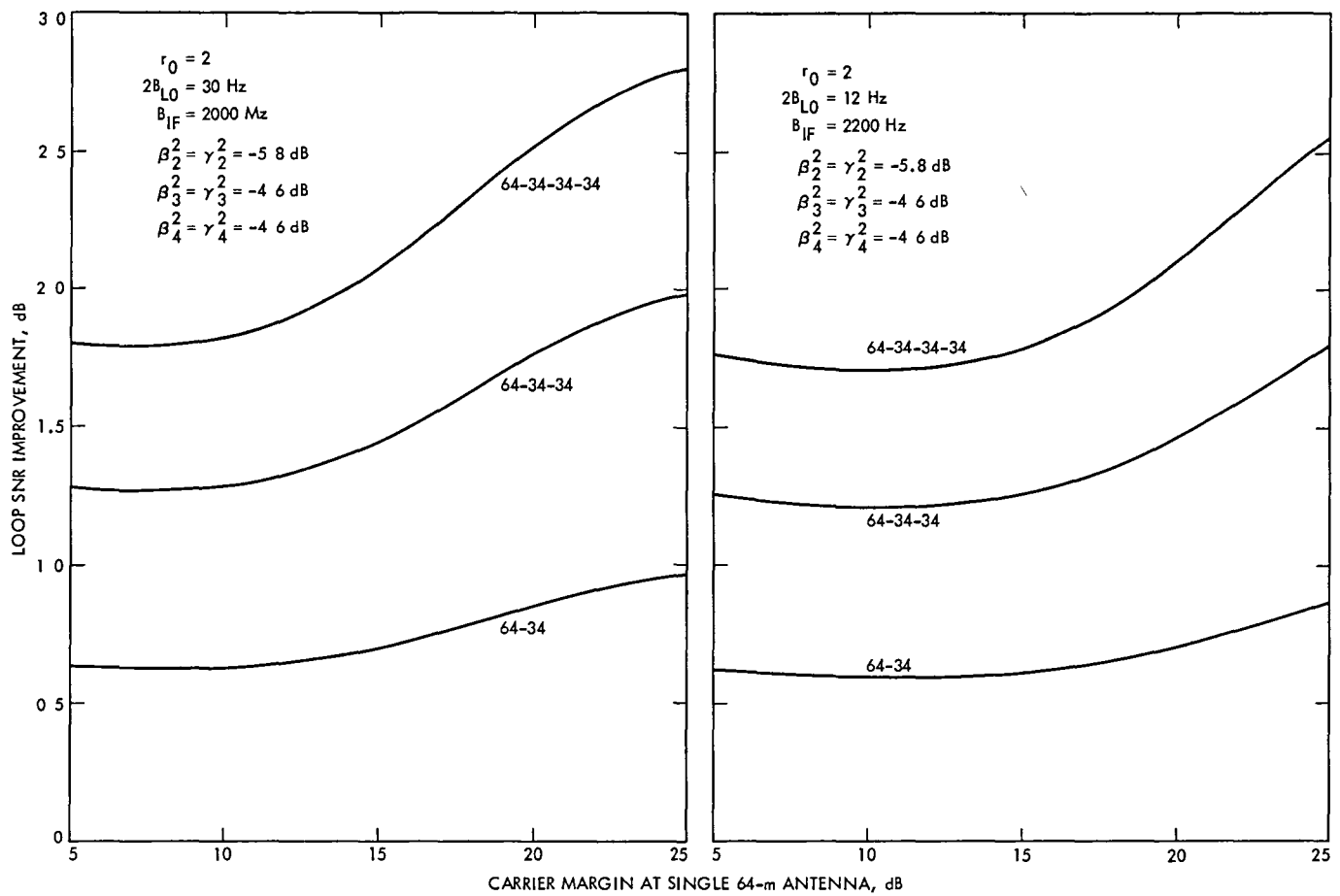


Fig 4 Loop SNR improvement for carrier arraying with respect to a single 64-m antenna station

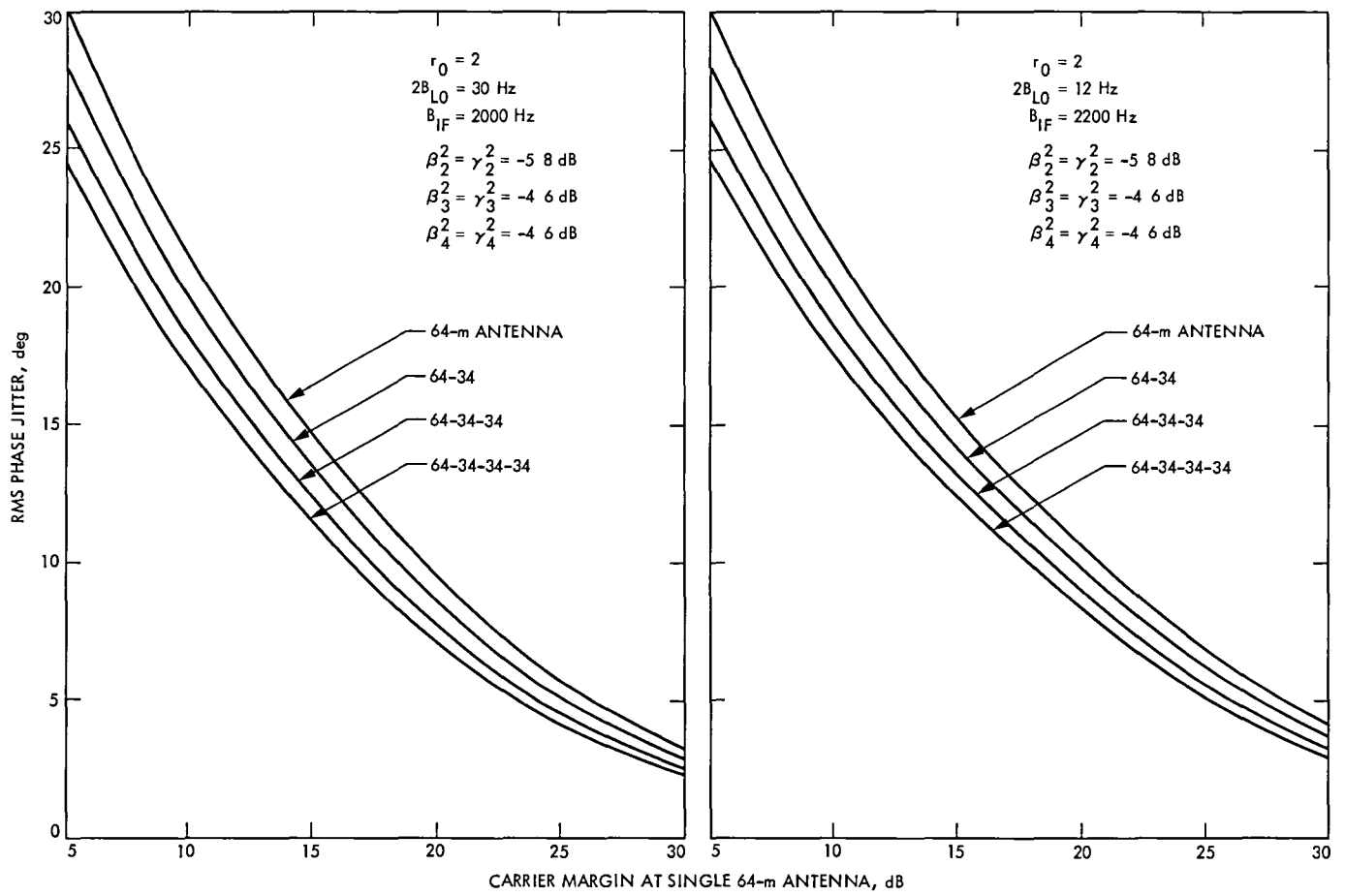


Fig 5 RMS phase jitter vs carrier margin for carrier arraying

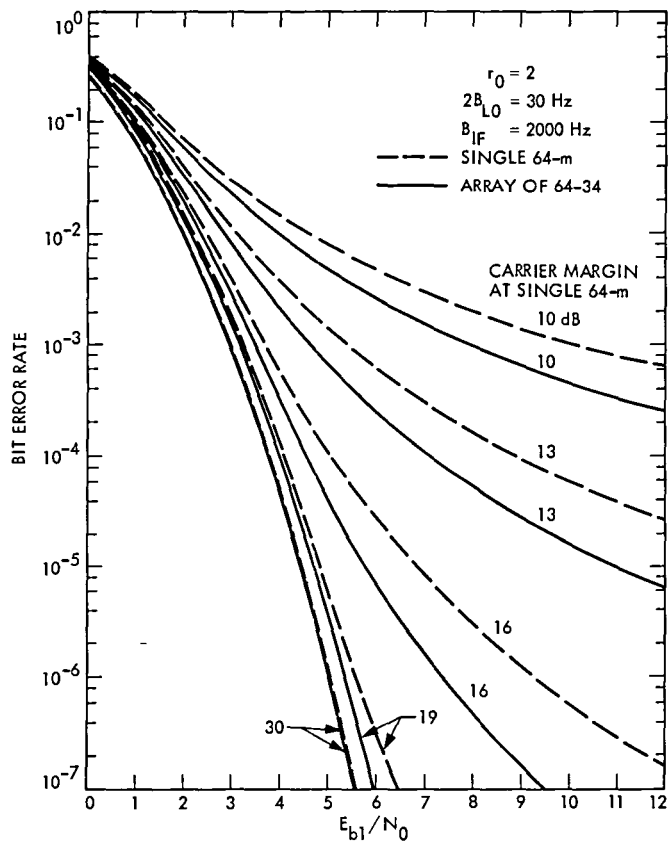


Fig. 6 Comparison of bit error rates of carrier arraying (64-34) and single 64-m, for various carrier margins at a single 64-m antenna

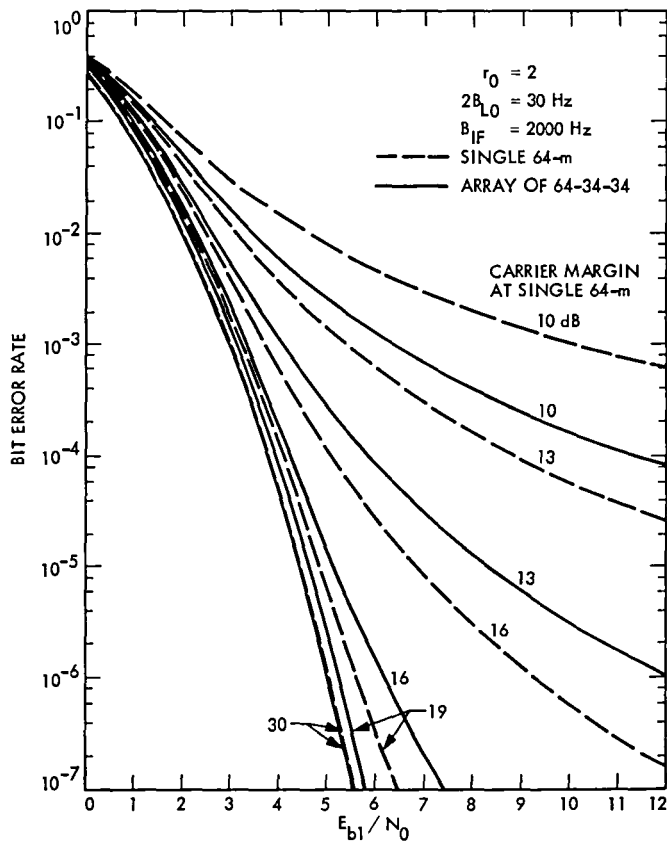


Fig. 7 Comparison of bit error rates of carrier arraying (63-34-34) and single 64-m, for various carrier margins at a single 64-m antenna

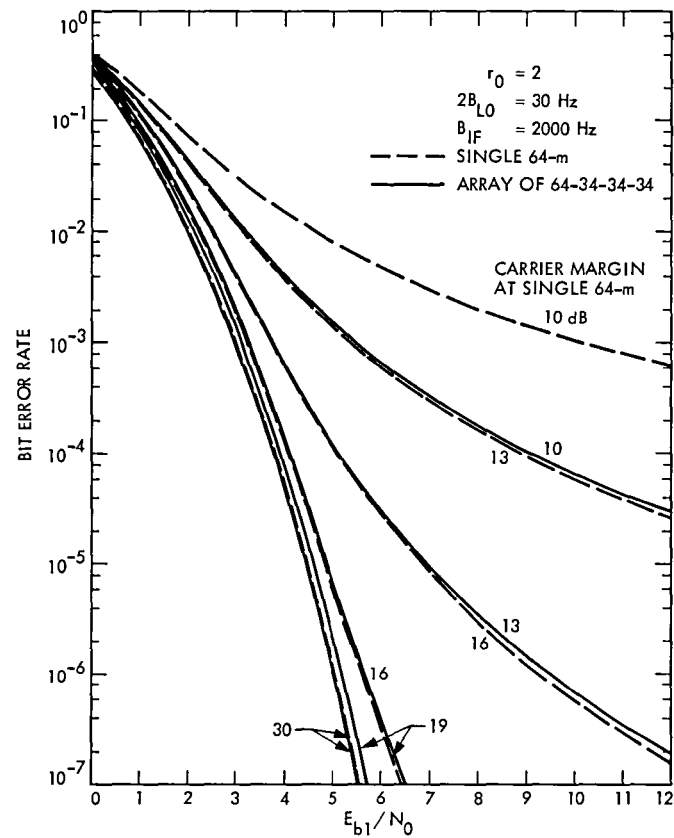


Fig 8 Comparison of bit error rates of carrier arraying (64-34-34-34) and single 64-m for various carrier margins at a single 64-m antenna

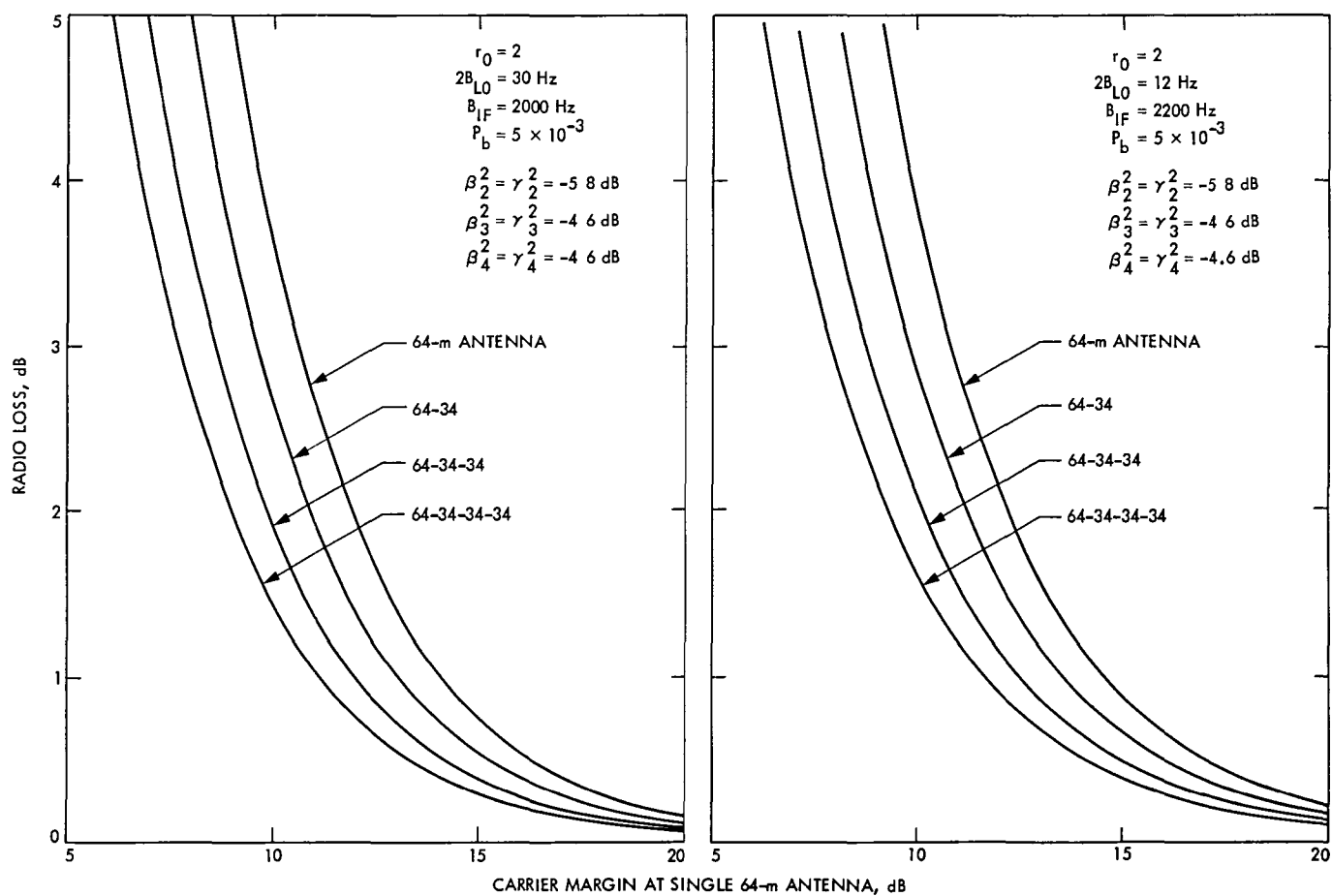


Fig 9 Radio loss for carrier arraying (Viterbi decoder, $K = 7$, $r = 1/2$, $Q = 3$)

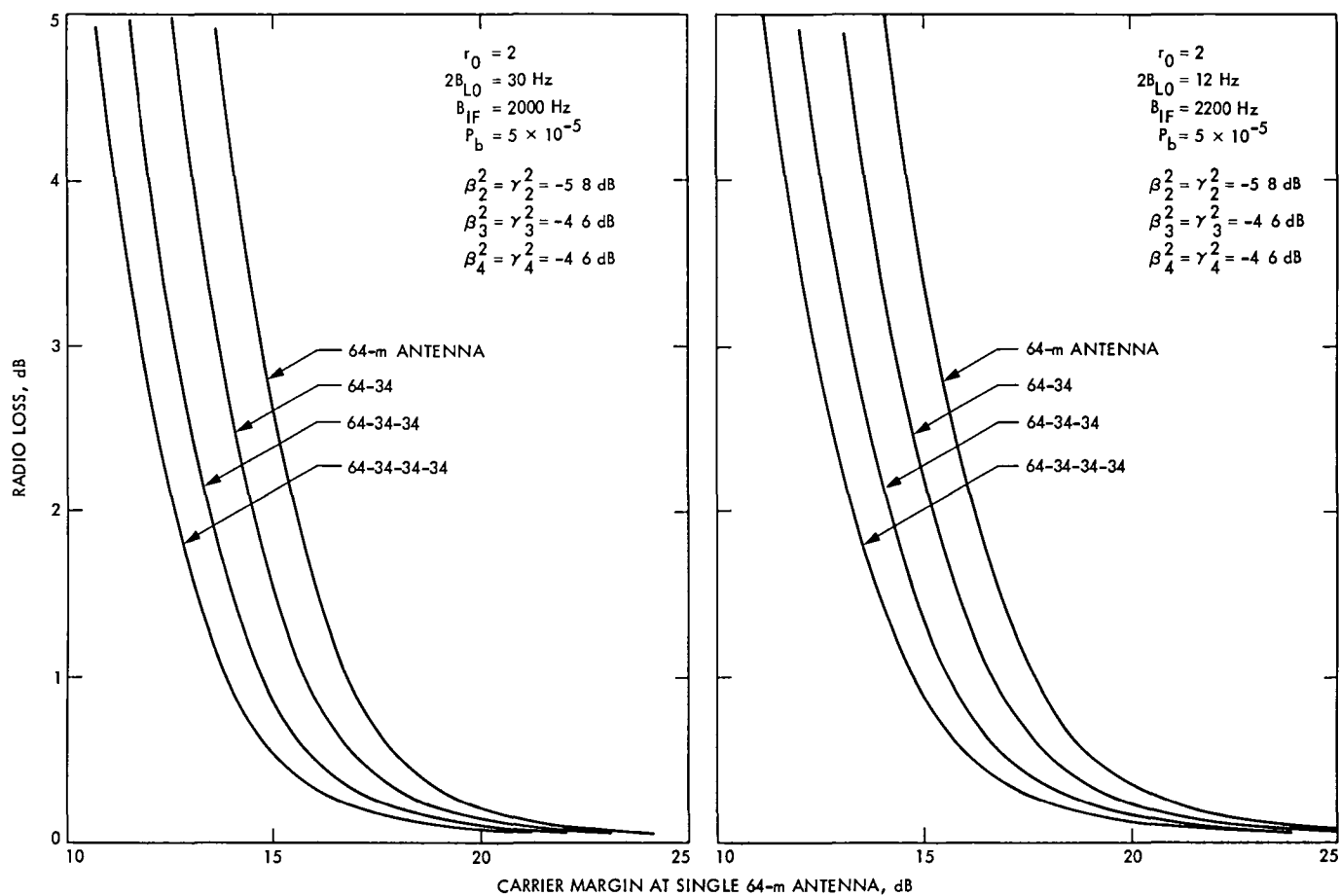


Fig. 10. Radio loss for carrier arraying (Viterbi decoder $K = 7$, $r = 1/2$, $Q = 3$)

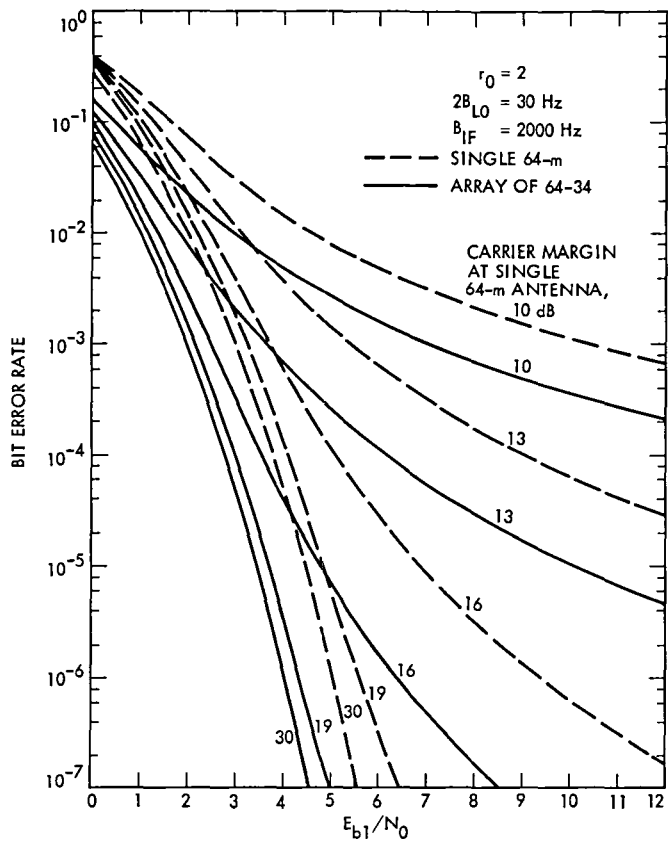


Fig 11 Comparison of bit error rates of combined carrier and baseband arraying (64-34) with a single 64-m, for various carrier margins at a single 64-m antenna

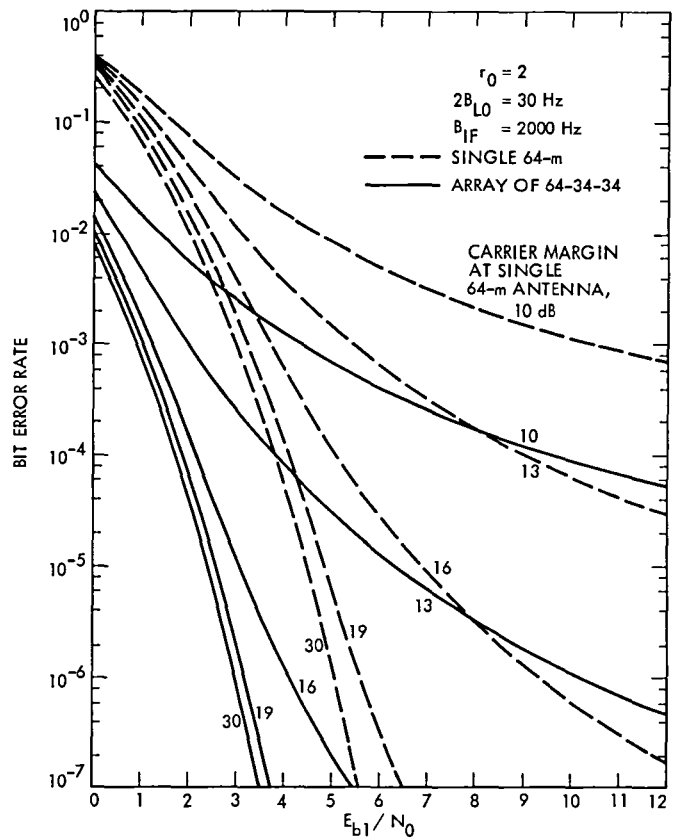


Fig 12. Comparison of bit error rates of combined carrier and baseband arraying (64-34-34) with a single 64-m, for various carrier margins at a single 64-m antenna

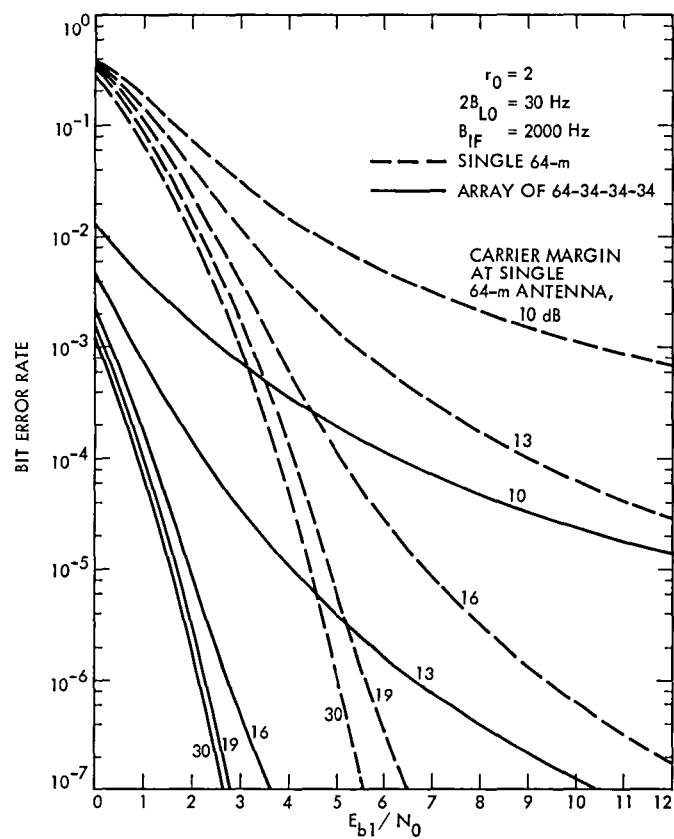


Fig 13. Comparison of bit error rates of combined carrier and baseband arraying (64-34-34-34) with a single 64-m, for various carrier margins at a single 64-m antenna

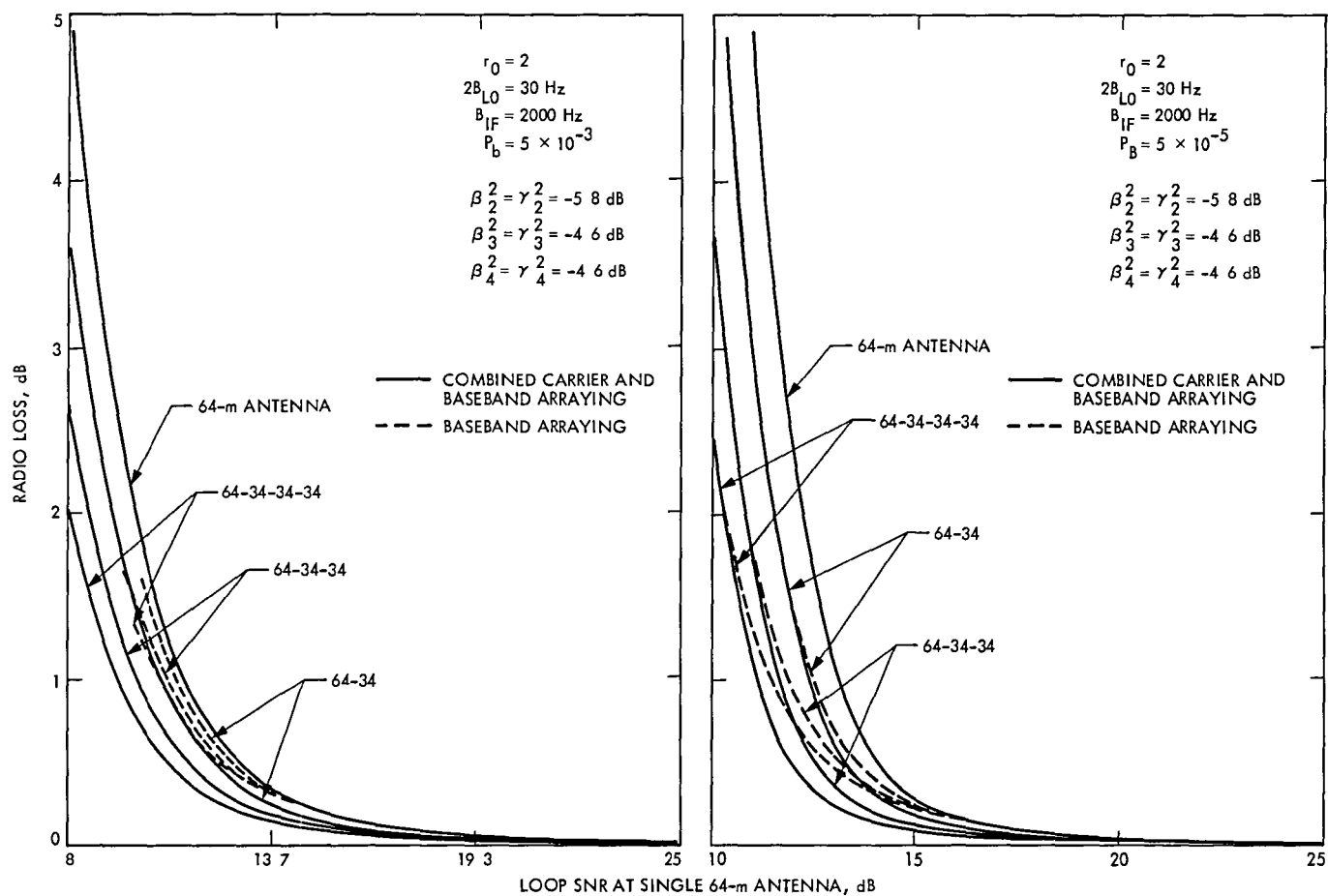


Fig 14. Comparison of radio loss for combined carrier and baseband arraying, and baseband arraying alone

Appendix A

Derivation of Phase Jitter

Introducing the Heaviside operator $p \triangleq d/dt$ from Fig 1 after passing $\tilde{S}(t)$ through the loop filter $F_1(s)$, the estimated carrier phase at the output of the VCO is

$$\begin{aligned}\hat{\theta}_1 &= \frac{K_1 F_1(p)}{p} \tilde{S}(t) \\ &= \frac{K_1 F_1(p)}{p} \left[\sum_{i=1}^N \sqrt{P_i} \beta_i \cos \theta_m \sin(\phi_1 + \phi_{ei}) + \sum_{i=1}^N \beta_i N_i(t, \phi_1 + \phi_{ei}) \right]\end{aligned}\quad (\text{A-1})$$

If the loop is now linearized (assume that $\sin(\phi_1 + \phi_{ei}) \cong \phi_1 + \phi_{ei}$), then we get

$$\hat{\theta}_1 = \frac{K_1 F_1(p)}{p} \left[\sqrt{P_1} \cos \theta_m (\theta_1 - \hat{\theta}_1) + \sum_{i=2}^N \sqrt{P_i} \beta_i \cos \theta_m (\theta_i - \hat{\theta}_i - \hat{\theta}_1) + \sum_{i=1}^N \beta_i N_i \right] \quad (\text{A-2})$$

where K_1 is total gain in the loop

Similarly for station i , $i = 2, 3, \dots, N$, we get

$$\hat{\theta}_i = \frac{K_i F_i(p)}{p} \left[\sqrt{P_i} \cos \theta_m \sin(\phi_1 + \phi_{ei}) + N_i(t, \phi_1 + \phi_{ei}) \right] \quad (\text{A-3})$$

If the loop is linearized, we get

$$\hat{\theta}_i = \frac{K_i F_i(p)}{p} \left[\sqrt{P_i} \cos \theta_m (\theta_i - \hat{\theta}_i - \hat{\theta}_1) + N_i \right] \quad (\text{A-4})$$

where K_i is total gain in the loop

Equation (A-4) can be solved for $\hat{\theta}_i$ as

$$\hat{\theta}_i = H_i(p) \left[(\theta_i - \hat{\theta}_1) + \frac{N_i}{\sqrt{P_i} \cos \theta_m} \right] \quad (\text{A-5})$$

where $H_i(p)$ is a closed-loop transfer function defined by (13) Substituting (A-5) in (A-2) we get

$$\hat{\theta}_1 = H_1(p) \left\{ \theta_1 + \frac{N_1}{\sqrt{P_1} \cos \theta_m} + \sum_{i=2}^N \beta_i \gamma_i [1 - H_i(p)] \left[(\theta_i - \hat{\theta}_1) + \frac{N_i}{\sqrt{P_i} \cos \theta_m} \right] \right\} \quad (\text{A-6})$$

Figure 2 can be realized from (A-5) and (A-6) Solving (A-6) for $\hat{\theta}_1$ we get

$$\hat{\theta}_1 = \frac{H_1(p) \left\{ \theta_1 + \frac{N_1}{\sqrt{P_1} \cos \theta_m} + \sum_{i=2}^N \beta_i \gamma_i [1 - H_i(p)] \left(\theta_i + \frac{N_i}{\sqrt{P_i} \cos \theta_m} \right) \right\}}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(p) [1 - H_i(p)]} \quad (\text{A-7})$$

Noting that $\phi_1 = \theta_1 - \hat{\theta}_1$ we get

$$\sigma_{\phi_1}^2 = E [\theta_1 - \hat{\theta}_1 - E(\theta_1 - \hat{\theta}_1)]^2 \quad (\text{A-8})$$

Using (A-7) in (A-8) we can get Eq (16)

$$\begin{aligned} \sigma_{\phi_1}^2 = & \frac{1}{2\pi\eta} \int \left| \frac{H_1(s)}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{01}}{2 P_1 \cos^2 \theta_m} \\ & + \sum_{i=2}^N \beta_i^2 \gamma_i^2 \frac{1}{2\pi\eta} \int \left| \frac{H_1(s) [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{0i}}{2 P_i \cos^2 \theta_m} \end{aligned} \quad (16)$$

Using assumption (18)

$$H_i(s) = H_2(s) \quad i = 3, 4, \dots, N \quad (18)$$

and substituting $H_2(s)$ given by (13) in (16) we get

$$\begin{aligned} \sigma_{\phi_1}^2 = & \frac{N_{01}}{2 P_1 \cos^2 \theta_m} \frac{1}{2\pi\eta} \int |1 + (\tau_1 + \tau_2)s + (\tau_1 \tau_2 + \tau_2^2/r_2)s^2 + \tau_1 \tau_2^2 s^3 / r_2| / \Delta|^2 ds \\ & + \sum_{i=2}^N \beta_i^2 \gamma_i^2 \frac{N_{0i}}{2 P_i \cos^2 \theta_m} \frac{1}{2\pi\eta} \int |(\tau_2^2 s^2 / r_2 + \tau_1 \tau_2^2 s^3 / r_2) / \Delta|^2 ds \end{aligned} \quad (\text{A-9})$$

$$\tau_1 \triangleq \tau_{21}$$

$$\tau_2 \triangleq \tau_{22}$$

where

$$\Delta = 1 + (\tau_1 + \tau_2)s + (\tau_1^2/r_1 + \tau_1 \tau_2 + G \tau_2^2/r_2)s^2 + (\tau_1^2 \tau_2/r_1 + G \tau_1 \tau_2^2/r_2)s^3 + \tau_1^2 \tau_2^2 s^4 / r_1 r_2$$

Using residue theorem we can evaluate the integrals. Assuming the ratio of the loop bandwidth of the PLL in station $i \neq 1$ is much narrower than the loop bandwidth of the PLL in station 1, then we can ignore $(B_{Li}/B_{L1})^n$ for $n \geq 2$ and we get the formula for $\sigma_{\phi_1}^2$ given by (19) as

$$\sigma_{\phi_1}^2 = \frac{r_1 B_{L1}}{(r_1 + 1) P_1 \cos^2 \theta_m} \left[\frac{1 + r_1 G + \xi(r_1 r_2 + r_2 G + (G - 1)^2)}{r_1 G^2 + r_2 \xi(r_1 G + G - 1)} N_{01} + \frac{1 + r_1 G + \xi(r_1 r_2 + 1)}{r_1 G^2 + r_2 \xi(r_1 G + G - 1)} \sum_{i=2}^N \beta_i^2 N_{0i} \right] \quad (19)$$

where ξ and G are given by (20) and (21)

Now we should evaluate $\sigma_{\phi_{ei}}^2$. Substituting (A-7) in (A-5) we get

$$\hat{\theta}_i = H_i(p) \left\{ \theta_i - \frac{H_1(p) \left[\theta_1 + \frac{N_1}{\sqrt{P_1} \cos \theta_m} + \sum_{j=2}^N \beta_j \gamma_j (1 - H_j(p)) \left(\theta_j + \frac{N_j}{\sqrt{P_j} \cos \theta_m} \right) \right]}{1 + \sum_{j=2}^N \beta_j \gamma_j H_1(p) [1 - H_j(p)]} + \frac{N_i}{\sqrt{P_i} \cos \theta_m} \right\} \quad (A-10)$$

Noting

$$\phi_{ei} = \theta_i - \theta_1 - \hat{\theta}_i$$

then

$$\sigma_{\phi_{ei}}^2 = E[\theta_i - \theta_1 - \hat{\theta}_i - E(\theta_i - \theta_1 - \hat{\theta}_i)]^2 \quad (A-11)$$

Substituting (A-10) in (A-11) we get Eq (46)

$$\begin{aligned} \sigma_{\phi_{ek}}^2 &= \frac{1}{2\pi j} \int \left| \frac{H_k(s) H_1(s)}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{01}}{2P_1 \cos^2 \theta_m} \\ &+ \sum_{\substack{i=2 \\ i \neq k}}^N \beta_i^2 \gamma_i^2 \frac{1}{2\pi j} \int \left| \frac{H_k(s) H_1(s) [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{0i}}{2P_i \cos^2 \theta_m} \\ &+ \frac{1}{2\pi j} \int \left| \frac{H_k(s) + H_k(s) H_1(s) \sum_{\substack{i=2 \\ i \neq k}}^N \beta_i \gamma_i [1 - H_i(s)]}{1 + \sum_{i=2}^N \beta_i \gamma_i H_1(s) [1 - H_i(s)]} \right|^2 ds \frac{N_{0k}}{2P_k \cos^2 \theta_m} \end{aligned} \quad (46)$$

Substituting $H_i(s)$ given by (13) in Eq (46), using assumption (18) and then applying the residue theorem, we can evaluate the integrals in Eq (46). Furthermore for simplicity of formula, since we have assumed

$$\frac{B_{Li}}{B_{L1}} \ll 1 \quad i \neq 1$$

then after ignoring all terms with $(B_{Li}/B_{L1})^n$ for $n \geq 1$, we can get Eq (48)

$$\begin{aligned} \sigma_{\phi_{ek}}^2 = & \left(\frac{G+r_2}{1+r_2} \right) \frac{N_{01} B_{L2}}{GP_1 \cos^2 \theta_m} + \frac{2r_2^2 + (1+G)r_2 + 2G}{(1+r_2)[2(1+G)r_2 + (1+G)^2]} \sum_{i=2}^N \beta_i^2 \frac{N_{0i} B_{L2}}{GP_1 \cos^2 \theta_m} \\ & + \left[1 - 2\beta_k \gamma_k \frac{(1+r_2)G + (1+2r_2)(r_2 - 1)}{(1+r_2)[2(1+G)r_2 + (1-G)^2]} \right] \frac{N_{0k} B_{L2}}{P_k \cos^2 \theta_m} \end{aligned} \quad (48)$$

Implications of the Putnam Software Equation on Confidence Levels for Project Success

R C Tausworthe
DSN Data Systems Section

This article investigates the implications of assumed powerlaw relationships among size, duration, and effort on the probability that a given software project will be completed within its estimated schedule and manpower resources. Specifically, software development tasks are treated as sample points in a probability space characterized by three random variables: size, duration, and resource expenditure. The completion confidence factor is then computed. The most astonishing conclusion is the low confidence factor of the average project, significantly less than 25%. This low confidence factor is the result of correlation of the project duration and work effort by a tradeoff relationship referred to as "Putnam's software equation."

I. Introduction

The early-on estimation of the resources and schedule required for the development and maintenance of software has resulted in several resource and schedule models that accept such inputs as the enormity of the task, the physical, environmental, human, and management constraints assumed or known to be in effect, the history base of similar and dissimilar experience, and the means, alternatives, and technology available to the task. Such models attempt to predict the performance of humans doing the work, the events and expenditures in the development process, and the characteristics of the resulting product. The goal of such models has been primarily to predict the average characteristics of an envisioned project in such a way as to suffice for planning purposes.

The prediction of human task group behavior, however, may be viewed as a problem in estimating events in a stochas-

tic process governed by an unknown, and probably nonstationary, probability function. An optimum model can predict only to the limit imposed by the statistical characterization of the human activity.

In addition to the mean behavior of software projects, the statistical treatment can also lead to estimators for project parameters that permit the evaluation of certain risk factors.

The optimal cost prediction model, however, would require the precise quantification of all technical, environmental, and human-behavior parameters, and would combine these into a mathematical formula producing maximal likelihood or minimal variance results. Lacking this precise quantification, the best that one may hope for in a cost model is that it accommodate the principal factors affecting the estimate variance (or project risk).

There are a number of software cost models in existence, fourteen of which are summarized in Ref. 1. Most of these are least-square-error fits using power-law relationships among size, resource, and schedule parameters

An IBM study (Ref. 2) reported the analysis of 60 software projects with respect to 68 situational variables believed to influence productivity. Of these, 29 showed a significant, high correlation with productivity, and were included in their estimation model. This model utilized power-law relationships for both the effects of the situational variables as well as the size, resource, and schedule variables

Nine such models were evaluated for Air Force use as reported in Ref. 3. Model accuracy was measured and found to be best whenever a particular model was calibrated using representative historical data. Calibration was found to have greater effect on estimating accuracy than the precise model form. Conclusions based on statistics from the University of Maryland (Ref. 4) tended to confirm this hypothesis on power-law models

All of the models above tended to focus on fitting measured statistics to relationships among pairs of parameters: effort vs size, duration vs size, effort vs duration, and average staff vs effort. Another model, the Rayleigh-Norden-Putnam model (Refs. 5 and 6), presupposes a power-law model among size, effort, and duration, a model calibrated using available industry data, however, the trivariate data for this calibration has not been published to the author's knowledge

This article investigates the implications of assumed power-law relationships among size, duration, and effort on the probability that a given software project will be completed within its estimated schedule and manpower resources. Specifically, the paper treats software development tasks as sample points in a probability space characterized by three random variables: size, measured in delivered lines of source code, duration, measured in months, and resource expenditure, measured in man-months. Power-law equations are used to describe the relationships between expected values of the software random variables, and log-normal probability functions are used to approximate marginal and conditional distributions

The paper then computes the *completion confidence factor*, defined as

$$C(T, W) = P\{t \leq T, w \leq W\}$$

i.e., the probability that the project duration, t , and work effort, w , will not simultaneously exceed values W and T , respectively.

II. Software Project Parameter Relationships

We shall suppose, for the purposes of this article, that a software development can be characterized by its final delivered size, the total effort expended, and the overall length of time required. We shall denote these quantities as

L = the number of kiloLines of delivered source code

w = the work effort required in man-months

t = the time duration in months

However, we shall find it more convenient to work with the logarithms of these

$$\log(L)$$

$$\log(w)$$

$$\log(t)$$

In this way, power-law relationships among L , w , and t become linear relationships in their logarithms

For a given program to be written, an infinite ensemble of projects would not all produce the same values of t , w , and L . Rather, one would observe a statistical distribution over the three-dimensional space spanned by these parameters. We shall thus treat software development characteristics as a probability space characterized by random variables $\log(L)$, $\log(w)$, and $\log(t)$, governed by a probability density function,

$$p(\log(t), \log(w), \log(L))$$

In terms of the usual expected-value operation, $E\{\cdot\}$, the point

$$(\tau_0, \omega_0, \lambda_0) = (E\{\log(t)\}, E\{\log(w)\}, E\{\log(L)\})$$

represents the characteristics of the *average project* across the hypothetical ensemble of projects writing the particular hypothetical program

A. Effort-Duration Tradeoff Characteristics

The average-project point $(\tau_0, \omega_0, \lambda_0)$ may be assumed to conform to a tradeoff law of the form

$$\tau_0 = g(\omega_0, \lambda_0)$$

That is, the average time required is influenced by the average effort applied and the average size of the task

Let us therefore define zero-mean random variables.

$$\tau = \log(t) - \tau_0$$

$$\omega = \log(w) - \omega_0$$

$$\lambda = \log(L) - \lambda_0$$

having standard deviations σ_τ , σ_ω , and σ_λ , respectively, and probability density function $p(\tau, \omega, \lambda)$. We seek to approximate the probability (confidence) that a project will be completed within schedule *and* manpower resources,

$$\begin{aligned} C(T, W) &= P\{t \leq T, w \leq W\} \\ &= P\{\tau \leq X, \omega \leq Y\} \\ &= \int_{-\infty}^X \int_{-\infty}^Y \int_{-\infty}^{+\infty} p(\tau, \omega, \lambda) d\lambda d\omega d\tau \\ &= \int_{-\infty}^X \int_{-\infty}^Y p(\tau, \omega) d\omega d\tau \end{aligned}$$

where

$$X = \log(T) - \tau_0$$

$$Y = \log(W) - \omega_0$$

The values X and Y are logarithmic schedule and effort margins measured from the average-project point. X and Y are related via the tradeoff law as follows

$$X = \log(T) - g(\log(W) - Y, \lambda_0)$$

It is clear that $C(T, W)$ depends on the average-project point, i.e., on the way expected effort has been traded with expected duration for a certain expected size. We shall optimize this situation by choosing the tradeoff to maximize the confidence,

$$C_0(T, W) = \max_{\tau_0 = g(\omega_0, \lambda_0)} C(T, W)$$

The condition for maximization of confidence is

$$\frac{dC(T, W)}{d\omega_0} = 0$$

which yields

$$\frac{dX}{dY} \int_{-\infty}^Y p(X, \omega) d\omega + \int_{-\infty}^X p(\tau, Y) d\tau = 0$$

Note that X may be eliminated from this equation by using the relationship imposed by the tradeoff law above. Thence the equation can be solved for Y , and then X and (τ_0, ω_0) found. These margin values will be denoted X_0 and Y_0 , respectively.

Note also that the solution above is the same as if $C(T, W)$ were maximized with respect to Y , subject to the constraint imposed by the tradeoff law

B. Putnam's Software Equation

Putnam (Ref. 6) has postulated a tradeoff between size, effort, and duration, and has given it the form

$$L = c_k w^p t^q$$

or

$$\log(L) = \eta + p \log(w) + q \log(t)$$

or, if we presume that the tradeoff is to be valid in a neighborhood of the average-project point,

$$\lambda = p\omega + q\tau$$

with

$$\lambda_0 = \eta + p\omega_0 + q\tau_0$$

The function $g(\)$ of the previous section is thus linear

$$\tau_0 = g(\omega_0, \lambda_0) = -\eta/q + \lambda_0/q - p\omega_0/q$$

and X and Y are related by

$$qX + pY = \eta + p \log(W) + q \log(T) - \lambda_0 = \Delta$$

The parameter Δ defined by this equation may be viewed as a *size margin*, because $\Omega = \eta + p \log(W) + q \log(T)$ is the virtual expected size when duration T and effort W are expended, and λ_0 is the expected size of the project. Note it is not necessary that the tradeoff equation apply at (T, W) . Ω is merely the constant defined by T and W in the tradeoff rule above

Putnam calls the tradeoff law the "software equation," and specifically postulates the values

$$p = 1/3$$

$$q = 4/3$$

for a *tradeoff exponent ratio* value of $r = q/p = 4$. These values were obtained by Putnam after a study of several large software implementation projects.

The author has been unable so far to locate collaborative published statistics in the open literature supporting the $r = 4$ figure. However, in private communications, the author has been informed by individuals having access to software productivity data bases that this particular statistic would not be difficult to compute, when undertaken.

The particular value $r = 4$ implies that a factor of 16 times as much manpower is required to shorten a schedule by a factor of 2. However, it well may be that r depends on project size. A value $r = 4$ may adequately describe localized changes in project parameters by a few percent about the average-project point for very large undertakings, but it may not apply to small-to-medium-sized efforts, where one may truly halve the overall duration by application of increased resources. For medium-scale efforts, it has been suggested that perhaps only an increase of 1.5 times the man-months effort is needed when the duration is halved, or $r = \log_2(1.5) = 0.484$. We shall discuss the influence of the value r on the confidence factor at greater length later in this article.

We shall therefore keep the values of the software equation coefficients undetermined for the present in order to be a bit more general in our approach, and later we shall solve for these, to the extent possible, using published empirical data.

C. Derivation of the Software Equation

If one were given a set of point-data $\{(\tau, \omega, \lambda)\}$ taken from an ensemble of projects of various durations, efforts, and sizes, then one could perform a least-squared-error fit to the data to determine general relationships among the random variables. Specifically, for the duration variable τ the best-fit curve would be (Ref. 7)

$$\tau_{\omega, \lambda} = E\{\tau|\omega, \lambda\}$$

We shall assume, in keeping with the Putnam model, that this relationship is approximately a linear one

$$\tau_{\omega, \lambda} = (\lambda - p\omega)/q$$

or

$$\lambda = p\omega + q\tau_{\omega, \lambda}$$

for appropriately chosen p and q . We have purposely chosen these coefficients to correspond to those appearing in the Putnam equation for later discussion.

The orthogonality principle in probability theory (Ref. 7) states that the error between τ and $\tau_{\omega, \lambda}$ is uncorrelated with both ω and λ . This provides two equations

$$E\{\omega\lambda\} = p\sigma_{\omega}^2 + qE\{\omega\tau\}$$

$$\sigma_{\lambda}^2 = pE\{\omega\lambda\} + qE\{\lambda\tau\}$$

In a similar fashion, one may seek to find relationships among each pair of random variables by finding the least-squared-error functions

$$\omega_{\lambda} = E\{\omega|\lambda\}$$

$$\tau_{\lambda} = E\{\tau|\lambda\}$$

$$\tau_{\omega} = E\{\tau|\omega\}$$

Here there is strong evidence (Refs. 2, 4, and 8) that a linear approximation to each of these functions is valid, say

$$\omega_{\lambda} = E\{\omega|\lambda\} = \alpha\lambda$$

$$\tau_{\lambda} = E\{\tau|\lambda\} = \beta\lambda$$

$$\tau_{\omega} = E\{\tau|\omega\} = \gamma\omega$$

Figures 1, 2, and 3 show scatter diagrams in support of these hypotheses.

Application of the orthogonality principle to these expectations produces the relationships

$$E\{\omega\lambda\} = \alpha\sigma_{\lambda}^2$$

$$E\{\lambda\tau\} = \beta\sigma_{\lambda}^2$$

$$E\{\omega\tau\} = \gamma\sigma_{\omega}^2$$

which may be substituted into the two equations found earlier to obtain

$$\alpha p + \beta q = 1$$

$$\alpha p + \alpha \gamma q = \alpha^2 \sigma_{\lambda}^2 / \sigma_{\omega}^2$$

provided that neither σ_{λ} nor σ_{ω} is zero.

Normally, one could solve these two simultaneous equations for p and q directly, whereupon a comparison with the Putnam values would be immediate, in terms of published best-fit parameters. However, we expect that β is approximately equal to $\alpha\gamma$, for the following reason. Note that α is the best-fit coefficient relating λ to ω , and γ is the best-fit coefficient relating ω to τ . Hence their product should be the coefficient relating λ to τ , and β is the actual best-fit coefficient relating λ to τ . If equality were the case, the two equations above would be singular, so no unique solution for p and q would then result.

But it is not the case that $\beta = \alpha\gamma$. If it were, the righthand sides of the two equations would have to be equal, a condition that would require

$$\alpha^2 = \sigma_\omega^2 / \sigma_\lambda^2 = \alpha^2 + \sigma_{\omega|\lambda}^2 / \sigma_\lambda^2$$

(we derive this latter relationship later), leading to a contradiction.

It is the case, however, that β is near enough to $\alpha\gamma$ that small errors are greatly magnified in the solutions for p and q . For this reason, it will be necessary to find p and q by other means, such as an actual two-parameter linear least-squares-fit of the raw data.

We can nevertheless express the values of p and q in terms of $r = q/p$,

$$p = \frac{1}{\alpha + \gamma r}$$

$$q = \frac{r}{\alpha + \gamma r}$$

In summary, we have assumed thus far only the following (1) that the function $E\{\tau|\omega, \lambda\}$, which minimizes the least-squared-error, is linear in both ω and τ , and (2) that each of the best-fit functions $E\{\omega|\lambda\}$, $E\{\tau|\lambda\}$, and $E\{\tau|\omega\}$ are also linear in their variables. These assumptions are all certainly approximately true, as supported by published analyses.

D. Program Size Statistics

For a given program to be written, one may estimate the mean and variance in size of the program by a number of techniques. Let us suppose that the method chosen produces representative estimates, and let us approximate the distribution of actual lines of code finally produced by a normal probability function,

$$p(\lambda) = Z(\lambda/\sigma_\lambda)$$

where (Ref. 9)

$$Z(x) = (2\pi)^{-0.5} \exp(-x^2/2)$$

Program size statistics are often postulated to fit the beta distribution rather than the normal, so as to avoid the theoretical occurrence of extremely large values of λ , which are permitted in the normal model (with low probabilities), but are absent in actuality. We shall discuss the implications of this later in the section on accuracy considerations.

E. Joint Effort-Duration Statistics

Let us now compute the function $p(\tau, \omega)$, which will permit us to evaluate the confidence integral for a particular project ensemble.

We shall assume that, if ω and λ are given, the density of the remaining random variable, τ , can be approximated by a normal density, in which the conditional mean in τ is determined by the software equation. Similarly, given λ , the density of ω can be estimated by the normal density, with its conditional mean as that value determined earlier. That is, we assume

$$p(\tau|\omega, \lambda) = Z((\tau - \tau_{\omega, \lambda})/\sigma_{\tau|\omega, \lambda})$$

$$p(\omega|\lambda) = Z((\omega - \omega_\lambda)/\sigma_{\omega|\lambda})$$

where $\sigma_{\tau|\omega, \lambda}$ is the standard deviation of τ given both ω and λ , and $\sigma_{\omega|\lambda}$ is the deviation of ω given λ . Scatter diagrams of ω vs λ published in the literature, such as that shown in Fig. 1, indicate that the $p(\omega|\lambda)$ approximation is reasonable. Further, such studies provide measured values for $\sigma_{\omega|\lambda}$.

The sought-for joint density is then

$$\begin{aligned} p(\tau, \omega) &= \int_{-\infty}^{+\infty} p(\tau|\omega, \lambda) p(\omega|\lambda) p(\lambda) d\lambda \\ &= \delta Z(\tau/\sigma_\tau) Z(\delta(\omega/\sigma_\omega - \rho\tau/\sigma_\tau)) \end{aligned}$$

This is a joint normal density in which the standard deviations σ_τ and σ_ω , correlation coefficient ρ , and parameter δ are given by

$$\sigma_\tau^2 = \sigma_{\tau|\omega, \lambda}^2 + \frac{\sigma_{\omega|\lambda}^2 (\sigma_{\omega|\lambda}^2 + \sigma_\lambda^2)}{r^2 \sigma_\omega^2}$$

$$\sigma_\omega^2 = \sigma_{\omega|\lambda}^2 + \alpha^2 \sigma_\lambda^2$$

$$\rho = -\sigma_{\omega|\lambda}^2 / r \sigma_{\tau} \sigma_{\omega}$$

$$\delta = (1 - \rho^2)^{-0.5}$$

(the equation above for σ_{ω}^2 is that referred to earlier, which yielded the contradiction in the equality $\beta = \alpha\gamma$)

Even though $\sigma_{\tau|\omega,\lambda}$ has not been estimated in the literature, the value of σ_{τ} can be evaluated, as will be shown later, to be

$$\sigma_{\tau}^2 = \sigma_{\tau|\lambda}^2 + \beta^2 \sigma_{\lambda}^2$$

All parameters needed to evaluate the joint density are therefore available from published data

Note, however, that since $|\rho| \leq 1$ and since r is unaffected by the value of σ_{λ} , the value of r is constrained by the lower bound

$$r \geq \frac{\sigma_{\omega|\lambda}}{\sigma_{\tau|\lambda}}$$

This relationship constrains p and q as follows

$$p \leq \frac{\sigma_{\tau|\lambda}}{\alpha \sigma_{\tau|\lambda} + \gamma \sigma_{\omega|\lambda}}$$

$$q \geq \frac{\sigma_{\omega|\lambda}}{\alpha \sigma_{\tau|\lambda} + \gamma \sigma_{\omega|\lambda}}$$

F. Effort Statistics

Integrating the joint density over τ yields the approximate marginal density of ω ,

$$p(\omega) = Z(\omega/\sigma_{\omega})$$

in which σ_{ω} is given above

G. Duration Statistics

Integration of the joint density above with respect to ω yields a normal approximation for $p(\tau)$ whose variance has a term $\sigma_{\tau|\omega,\lambda}^2$. However, the variance may be computed another way the published statistics of τ vs λ indicate that a normal approximation is appropriate

$$p(\tau|\lambda) = Z((\tau - \tau_{\lambda})/\sigma_{\tau|\lambda})$$

where $\sigma_{\tau|\lambda}$ is the standard deviation in τ , given λ , obtained from the published data. Averaging over λ produces the normal marginal density for τ ,

$$p(\tau) = Z(\tau/\sigma_{\tau})$$

in which the variance of τ is, as stated earlier,

$$\sigma_{\tau}^2 = \sigma_{\tau|\lambda}^2 + \beta^2 \sigma_{\lambda}^2$$

Thus, the deviation in duration can be calculated directly from the measured and estimated deviations in ω and λ , respectively. Hence, the need to compute $\sigma_{\tau|\omega,\lambda}^2$ is removed

H. Computation of Confidence Factor

Integration of the joint normal density $p(\tau, \omega)$ to produce $C(T, W)$ as indicated earlier yields a known (but untabulated) function (Ref 9)

$$C(T, W) = L(-X/\sigma_{\tau}, -Y/\sigma_{\omega}, \rho)$$

where $L(x, y, \rho)$ is defined as the double integral

$$L(x, y, \rho) = \int_x^{\infty} \int_y^{\infty} \delta Z(u) Z(\delta(v - \rho u)) du dv$$

Values for $L(x, y, \rho)$ may be found by numerical integration

One particular case of interest can be evaluated directly, however, namely the confidence factor for the average project, $C(T_0, W_0)$. This results when $X = Y = 0$, giving

$$\begin{aligned} C(T_0, W_0) &= L(0, 0, \rho) \\ &= \frac{1}{4} + \frac{\arcsin \rho}{2\pi} \end{aligned}$$

Note that since ρ is a negative quantity, the *confidence in average project success is less than 25%*. To raise the confidence factor to reasonable levels, it is therefore necessary to increase the schedule and effort margins, i.e., X and Y , to positive values

Evaluation of the previous conditions for finding X_0 and Y_0 produces the equation

$$Z(aX_0) P(b - cX_0) = h Z(d - eX_0) P(f + gX_0)$$

where $P(x) = (1 + \operatorname{erf}(x/2^{0.5}))/2$, in which $\operatorname{erf}(\cdot)$ is the well-known error function (Ref 9), and the coefficients a

through h are given by

$$a = 1/\sigma_\tau$$

$$b = d\delta$$

$$c = \delta(r/\sigma_\omega + \rho/\sigma_\tau)$$

$$d = \Delta/p\sigma_\omega$$

$$e = r/\sigma_\omega$$

$$f = -b\rho$$

$$g = c\rho + 1/\delta\sigma_\tau$$

$$h = e\sigma_\tau$$

The solution X_0 to this equation may be found numerically using Newton's method, and then the corresponding Y_0 computed

The confidence factor is then computed by evaluating

$$C_0(T, W) = L(-X_0/\sigma_\tau, -Y_0/\sigma_\omega, \rho)$$

It is interesting to note in the optimization equation when $\Delta = 0$ that $X_0 = Y_0 = 0$ is *not* a solution. Instead, one can show that $X_0 < 0$ and $Y_0 > 0$ for this situation. This means that the confidence will be somewhat greater for a project expected to last somewhat longer than $\log(T)$ with somewhat less manpower than $\log(W)$ than it will be for a project with expected duration equal to $\log(T)$ and expected manpower equal to $\log(W)$.

III. Evaluation

In this section we shall compute the confidence characteristics of projects taken from two sources of published data. These two sources were not chosen because of their similarities or differences, or because the data was particularly extensive or internally consistent. Rather, the sources were used because they contained enough information to compute the model parameters of interest.

A. Walston-Felix Data

Data published by Walston and Felix (Ref. 2) provide the following parameter values of interest

$$\alpha = 0.91$$

$$\sigma_{\omega|\lambda} = 0.92$$

$$\beta = 0.36$$

$$\sigma_{\tau|\lambda} = 0.542$$

$$\gamma = 0.35$$

$$\sigma_{\tau|\omega} = 0.419$$

$$\beta - \alpha\gamma = 0.0415$$

Even though (Ref. 2) contains a means for approximately normalizing the effects of the 29 noted situational and environmental variables, the values above do not reflect this (partial) normalization.

We shall assume, purely for illustrative purposes, that a PERT (Ref. 10) estimation scheme is used to estimate λ and its variance, and that $\lambda_{max} = \log(1.5\lambda_0)$ and $\lambda_{min} = \log(\lambda_0/2)$, so that we have

$$\sigma_\lambda = (\lambda_{max} - \lambda_{min})/6 = 0.183$$

We then obtain from the formulas above

$$\sigma_\omega = 0.935$$

$$\sigma_\tau = 0.546$$

The bounds on r , p , and q are

$$r \geq 1.7$$

$$p \leq 0.665$$

$$q \geq 1.13$$

A range of confidence factor calculations for the average-project point is shown in Table 1.

The Putnam value $p = 1/3$ would require $r = 5.97$, according to the formulas above, rather than the $r = 4$ value used by Putnam. The $r = 4$ value, on the other hand, together with the measured values above, produces $p = 0.433$. If $q = 4/3$ were the case, then r would have to be 2.28.

The average-project confidence factors for these cases are shown in Fig. 4 as a function of r . The optimized confidence factor characteristics are plotted in Fig. 5 as a function of the size margin parameter Δ . The optimum duration and effort margins are given in Fig. 6.

B. Freburger-Basili Data

Data published by Freburger and Basili (Ref. 4) give

$$\alpha = 0.986$$

$$\sigma_{\omega|\lambda} = 0.378$$

$$\beta = 0.203 \quad \sigma_{\tau|\lambda} = 0.315$$

$$\gamma = 0.210 \quad \sigma_{\tau|\omega} = 0.300$$

$$\gamma = 0.210$$

$$\beta - \alpha\gamma = -0.0004$$

The data from which these parameters were derived were taken from a more carefully controlled environment and situation than the Walston-Felix data was. Computations using these values result in

$$\sigma_{\omega} = 0.419$$

$$\sigma_{\tau} = 0.317$$

The bounds on r , p , and q are

$$r \geq 1.202$$

$$p \leq 0.807$$

$$q \geq 0.971$$

The values of the correlation coefficient and average-project confidence factor deriving from these values are shown in Table 2.

A value of $r = 4$ implies $p = 0.547$, rather than the Putnam $1/3$, and the $p = 1/3$ value implies $r = 9.59$. A q -value of $4/3$ requires $r = 1.82$.

The average-project confidence factors above are plotted in Fig. 4 vs r for comparison with the Walston-Felix data. The optimized confidence factor behavior as a function of the size margin is shown in Fig. 7, and the optimum duration and effort margins are given in Fig. 8.

C. Comparison of Results

The Freburger-Basili data generally show a higher confidence factor than do the Walston-Felix data, both for the average project and the optimized project. The difference may be due to the extreme care and consistency taken in recording the data in the former case. The latter data were taken from a diverse set of projects in a wide variation of environments and situational factors, whereas the former were taken under very controlled and recorded conditions. If both sets of data were normalized to remove situational effects, as described in the Walston-Felix paper, then perhaps the agreement would be closer.

Nothing can be learned about the value of r from either of these data sets, as published. The variation of X_0 and Y_0 for the two cases is radically different.

IV. Accuracy Considerations

Accuracy in the figures produced by the model above depends not only on the assumptions about normality of the logarithms and linearity of the conditional expectations, but also on the accuracy of the inferences from statistical measures to the model parameter values. By assuming an effort-duration tradeoff law, we have admitted that the average project facing a given situation can be influenced by the expected allocation of effort and duration. Yet the published statistics derive from an ensemble of projects tasked with an ensemble of very different situations. The marked differences between the Walston-Felix data and the Freburger-Basili data are strong indicators that situational factors, such as technology, organization, experience, and environment, do influence project performance. Walston and Felix confirm this hypothesis in their article, where they find significant correlation between performance and 29 situational variables.

Therefore, the published statistics represent an averaging over the model we have presented here with respect to the ensemble of industry projects. We may thus expect that the use of measured variances taken from such data will exceed the variances assumed in the model above. In this respect, use of the Walston-Felix or Freburger-Basili numbers may produce somewhat pessimistic results, unless the situational factors can be normalized to some extent by an appropriate method.

Frequently, the statistics of a given bounded data set are presumed to be characterized by a beta density,

$$p(x) = A (x - x_1)^b (x_2 - x)^c$$

over the range $x_1 \leq x \leq x_2$, with $p(x) = 0$ elsewhere. A is a constant chosen to make the density have unit area. Typically, values $b = c = 2$ are used in the so-called PERT (Ref. 10) estimation technique.

This density has peak value x_0 given by

$$x_0 = \frac{bx_2 + cx_1}{d}$$

where $d = b + c$. The mean and variance of the distribution are

$$\mu = \frac{x_1 + dx_0 + x_2}{d + 2}$$

$$\sigma^2 = \frac{(d + bc + 1)(x_2 - x_1)^2}{(d + 2)^2 (d + 3)}$$

The values of b and c for a given d may be adjusted to accommodate nonsymmetry in the distribution

Let us now consider how well the normal probability density approximates the beta density. Let us compare both as zero-mean, unit-variance densities with respect to two measures: first, the root-sum-square error, and second, the probability that the normal variate exceeds the beta-distribution cutoff point. The conditions for the beta density having zero mean and unit variance are

$$x_1 = -(b + 1) \left[\frac{d + 3}{bc + d + 1} \right]^{0.5}$$

$$x_2 = -\frac{(c + 1)}{(b + 1)} x_1$$

The first comparison function, the root-sum-error, is

$$\epsilon_1 = \left[\int_{-\infty}^{+\infty} (p(x) - Z(x))^2 dx \right]^{0.5}$$

which can be found by numerical integration. The second comparison function can be evaluated directly,

$$\begin{aligned} \epsilon_2 &= [1 - P(-x_1)] + [1 - P(x_2)] \\ &= 0.5 [\operatorname{erfc}(-x_1/2^{0.5}) + \operatorname{erfc}(x_2/2^{0.5})] \end{aligned}$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function (Ref. 9)

The values of both ϵ_1 and ϵ_2 are shown in Table 3 below. From these one may note that both errors decrease with d , and for a given d , both are least when b and c are equal (i.e., when the density is symmetric, as is $Z(x)$). We may therefore conclude that for a reasonable d (e.g., $d = 4$ in the PERT technique), the error introduced into the model by assuming a normal distribution, rather than a beta distribution, is negligible with respect to the much larger uncertainties involved with estimating the parameters required by the model.

V. Summary and Conclusions

There is something unsettling, or at least disappointing, in the results of these analyses, that the average project will probably fail in meeting one of its performance goals. How-

ever, this likelihood is implied by the combined assumptions of the popular software cost estimators assembled into the model of this article. All of these assumptions individually sound reasonable, fit together in a consistent, logical way, and are backed up by measurements of one sort or another.

Moreover, the model predicts that a significant contingency bias in planned manpower and schedule will be required to reach acceptable engineering risk levels.

What may be even more unsettling is that this theory may apply to other kinds of projects as well, where there is a log-linear, manpower-schedule tradeoff possible. A study of productivity and schedule statistics in this area might be very revealing.

However, for the moment, we may consider that part of what has been called "the software crisis," i.e., that most projects seem to fail one way or another, may not be the fault of either the programmer or his management, for even if they were able to estimate exactly what the average project would do in their given situation, the odds are that they would still fail, if they planned for the average.

What this paper shows, nevertheless, is that *planning for an average project contributes to the crisis*. To succeed within performance goals, it is necessary to do the following things:

- (1) Estimate carefully the size, and bounds on size, of the task.
- (2) Negotiate manpower and schedule constraints.
- (3) Determine the risks associated with failure of the project in both manpower and schedule dimensions.
- (4) Negotiate an appropriate confidence factor under which the risks are acceptable.
- (5) Determine (T, W) that will produce the desired confidence within manpower and schedule constraints using the model above.
- (6) Schedule the task to utilize manpower W and duration T .
- (7) If a (T, W) cannot be found that is compatible with risk and constraints on manpower and schedule, renegotiate.

As mentioned earlier, the model will probably produce somewhat pessimistic results if published industry parameter values are used. Accuracy in estimating confidence factors and planned performance margins can come only through precalibration of the model parameters in the particular software development environment and normalization of the situational

factors that may influence that particular project in the given environment. The Thibodeau study (Ref. 3) also confirmed that such calibration is needed for more basic modeling accuracy considerations

To promote better accuracy, it is necessary to develop a basic software cost model, such as that reported in Ref 2 and extended by the author in Ref 11, that will tend to normalize

the environmental and situational factors a particular project faces, to maintain an historical archive of measured project characteristics, and to analyze the collected data with respect to the model to extract the needed model parameters

In particular, the value of r needs to be explicitly measured. The analyses of this paper are inconsistent with the Putnam p , q , and q/p .

References

1. *Quantitative Software Models*, Report SRR-1, Data and Analysis Center for Software, RADC, Griffiss Air Force Base, New York, AFB, N.Y , March 1979
2. Walston, C E , and Felix, C P., "A Method of Programming Measurement and Estimation," *IBM System Journal*, Vol 16, No. 1, 1977
3. Thibodeau, R , *An Evaluation of Software Cost Estimating Models*, Report CR-1-940, prepared for Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, New York, by General Research Corp., Huntsville, Ala , April 10, 1981.
4. Freburger, K , and Basili, V R , *The Software Engineering Laboratory Relationship Equations*, TR-764, SEL-3, NSG-5123, University of Maryland, College Park, Md , May 1979
5. Norden, P. V , "Project Life Cycle Modeling," Software Life-Cycle Management Workshop, United States Army Computer Systems Command, August 1977, pp. 217-306
6. Putnam, L H , "Progress in Modeling the Software Life Cycle in a Phenomenological Way to Obtain Engineering Quality Estimates and Dynamic Control of the Process," in *Software Cost Estimating and Life-Cycle Control*, IEEE Computer Society Press, Los Alamitos, Ca , 1980, pp 183-206.
7. Papoulis, A , *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Book Co , N.Y , 1965, pp. 104, 147, 241-243, 250-258
8. Nelson, R , *Software Data Collection and Analysis, (Partial Report)*, Rome Air Force Development Center, Data and Analysis Center for Software, Griffiss Air Force Base, New York, September 1978
9. *Handbook of Mathematical Functions*, AMS-55, National Bureau of Standards, U S. Government Printing Office, Washington D.C , 1964, pp. 297-298, 931-937
10. The author is unable to trace the definitive reference for this work. The usage cited is exposed in Putnam, L H , "Example of an Early Sizing, Cost and Schedule Estimate for an Application Software System," *COMPSAC '79*, November 13-16, 1978
11. Tausworthe, R C , *Deep Space Network Software Cost Estimation Model*, Publication 81-7, Jet Propulsion Laboratory, Pasadena, Ca , Apr 15, 1981.

Table 1. Average project confidence factors as a function of the tradeoff ratio r , from Walston-Felix data

r	p	q	ρ	$C(T_0, W_0)$
2.00	0.621	1.24	-0.829	0.094
2.28	0.586	1.33	-0.729	0.120
4.00	0.433	1.73	-0.415	0.182
5.97	0.333	1.99	-0.277	0.205

Table 2. Average project confidence factors as a function of the tradeoff ratio r , from Freburger-Basil data

r	p	q	ρ	$C(T_0, W_0)$
1.83	0.730	1.33	-0.590	0.150
2.00	0.711	1.42	-0.539	0.159
4.00	0.547	2.19	-0.269	0.207
9.59	0.333	3.20	-0.112	0.232

Table 3. Beta density vs normal density goodness of fit

b	c	d	ϵ_1	ϵ_2
1	1	2	0.088	0.025
1	2	3	0.091	0.024
1	3	4	0.107	0.031
1	4	5	0.122	0.037
2	1	3	0.091	0.024
2	2	4	0.057	0.008
2	3	5	0.058	0.008
2	4	6	0.067	0.010
3	1	4	0.107	0.031
3	2	5	0.058	0.008
3	3	6	0.043	0.003
3	4	7	0.043	0.003
4	1	5	0.122	0.036
4	2	6	0.067	0.010
4	3	7	0.043	0.003
4	4	8	0.033	0.001

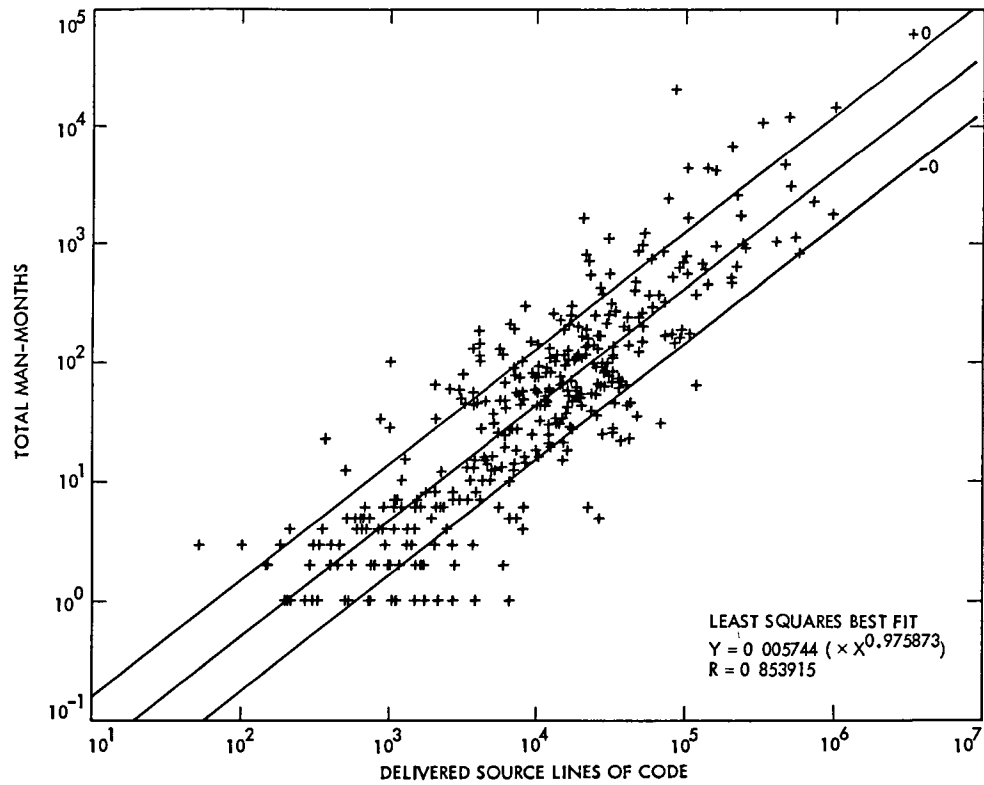


Fig. 1. RADC software project effort vs size data (Ref. 8)

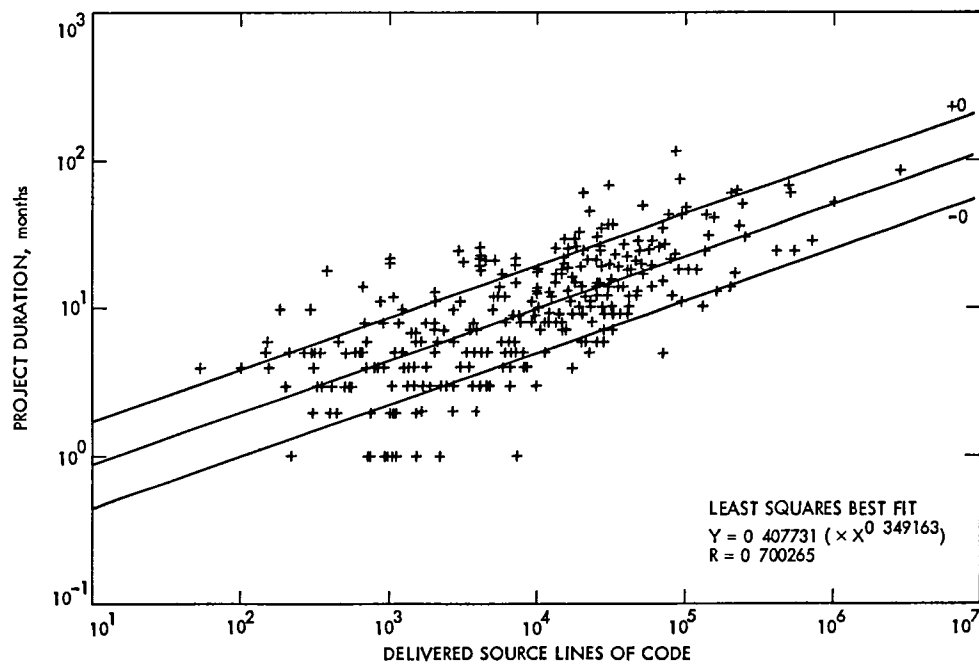


Fig. 2. RADC software project duration vs size data (Ref. 8)

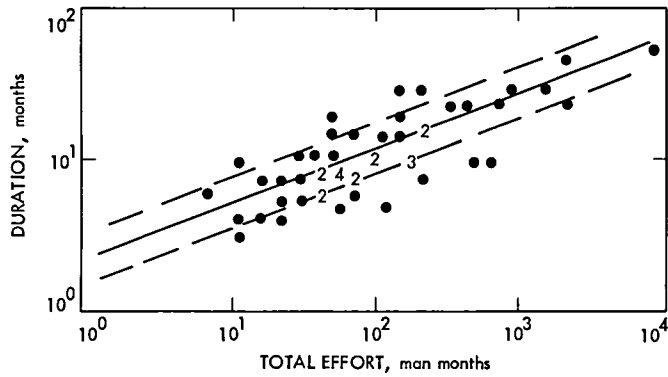


Fig 3 IBM software project duration vs effort data (Ref 2)

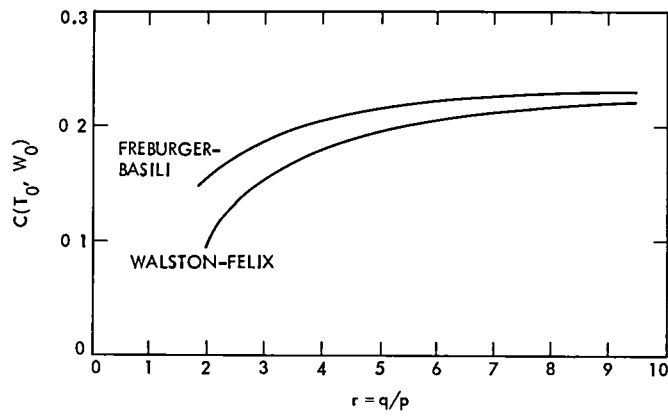


Fig 4 Average-project confidence factors as a function of r for both the Walston-Felix data and Freburger-Basili data

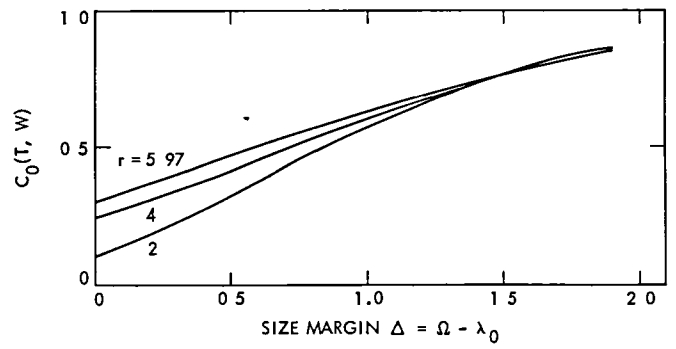


Fig. 5. Optimum project confidence factor for given size margin Δ , Walston-Felix data

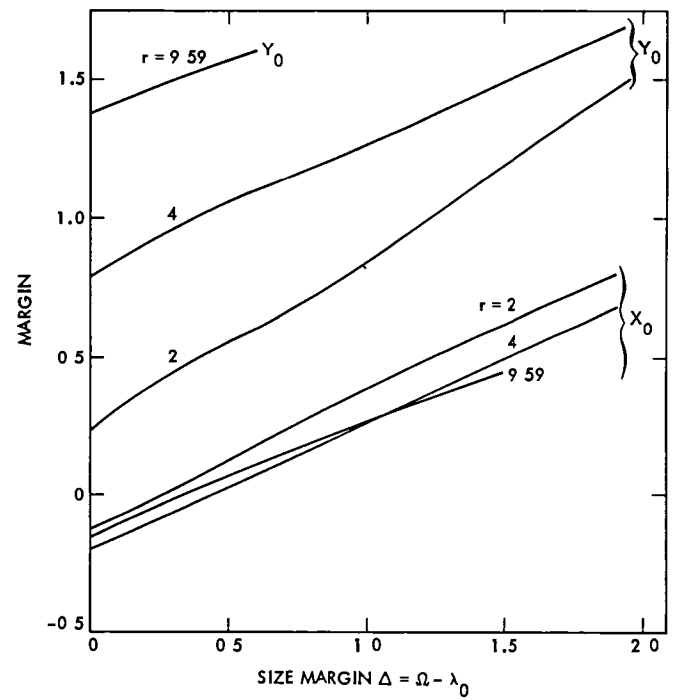


Fig. 6. Optimized duration and effort margins for given size margin Δ , Walston-Felix data

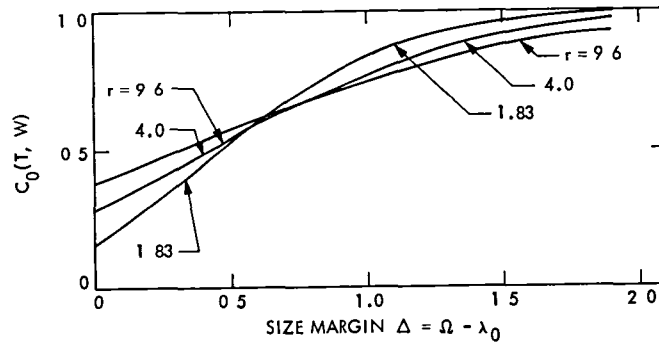


Fig. 7. Optimum project confidence factor for given size margin Δ , Freburger-Basili data

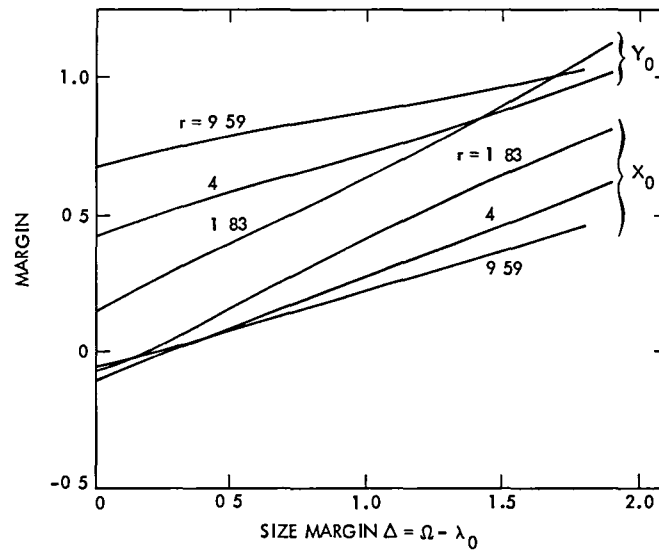


Fig. 8. Optimized duration and effort margins for given size margin Δ , Freburger-Basili data

Gain Uncertainty of Arrayed Antennas

B. L. Seidel

Radio Frequency and Microwave Subsystems Section

The gain uncertainty of an array of antennas is derived in terms of the gain uncertainty of the individual elements. For the case where the gain uncertainties of the individual elements are about equal, the gain uncertainty (uncorrelated errors) of the array is less than that of the individual elements. In the DSN, the gain uncertainty of an array composed of a 64-meter antenna and one or more 34-meter antennas is most sensitive to the uncertainty of the gain of the 64-meter antenna. For example, a 64-meter antenna ($G \cong 61.7 \pm 0.1$ dB) and a 34-meter antenna ($G_2 \cong 56.1 \pm 0.3$ dB) result in an array with $G \cong 62.8 \pm 0.1$ dB (uncorrelated), ± 0.15 dB (correlated). For this example, the array gain uncertainty is not significantly affected by the smaller element gain tolerance.

I. Introduction

When a 64-meter and one or more 34-meter DSN antennas are arrayed, the increase in the gain of the array relative to the gain of the 64-meter antenna is of the same order of magnitude (in dB) as the uncertainty in the gain of each of the antennas. With this in mind, it might appear that only minimal, if any, operational advantage can be gained by arraying, since conservative telecommunications link design requires one to consider the sum of the negative tolerances. This report shows that arraying is valuable to the link performance even under the conditions stated. It is recognized that many factors go into making up the performance of an antenna array. This analysis only addresses the performance uncertainty of the array due to the uncertainty in the gain of the individual antennas.

$$G = \sum_{i=1}^n G_i \quad (1)$$

and (assuming statistical independence) the uncorrelated array gain uncertainty is

$$\sigma_u = \left[\sum_{i=1}^n (\sigma_i)^2 \left(\frac{\partial G}{\partial G_i} \right)^2 \right]^{1/2} \quad (2)$$

but

$$\frac{\partial G}{\partial G_1} = \frac{\partial G}{\partial G_2} = \dots = \frac{\partial G}{\partial G_n} = 1$$

Therefore

$$\sigma_u = \left[\sum_{i=1}^n (\sigma_i)^2 \right]^{1/2} \quad (3)$$

II. Theory

Consider a set of antennas with gains G_1, G_2, \dots, G_n and uncertainties $\sigma_1, \sigma_2, \dots, \sigma_n$ respectively. The array gain is given by

and (assuming statistical dependence) the correlated array gain uncertainty is or

$$\sigma_c = \sum_{i=1}^n \sigma_i \quad (4)$$

where σ_u assumes that the σ_i s are uncorrelated and σ_c assumes that the σ_i s are fully correlated and add in the worst possible manner. Note that the above relationships hold when the gains and uncertainties are expressed as ratios. Since

$$G = 10^{G_{dB}/10} \quad (5)$$

and

$$\sigma = \sigma_{dB} \left(\frac{\partial G}{\partial G_{dB}} \right) \quad (6)$$

we get

$$\sigma = \frac{\ln 10}{10} (G) (\sigma_{dB}) \quad (7)$$

or

$$\sigma_{dB} = \frac{10}{\ln 10} \left(\frac{\sigma}{G} \right) \quad (8)$$

substituting (1) and (3) into (8) gives

$$\sigma_{u_{dB}} = \frac{10}{\ln 10} \frac{\left[\sum_{i=1}^n (\sigma_i)^2 \right]^{1/2}}{\sum_{i=1}^n G_i} \quad (9)$$

or

$$\sigma_{u_{dB}} = \frac{\left[\sum_{i=1}^n (G_i \sigma_{i_{dB}})^2 \right]^{1/2}}{\sum_{i=1}^n G_i} \quad (10)$$

while substituting (1) and (4) into (8) gives

$$\sigma_{c_{dB}} = \frac{10}{\ln 10} \frac{\sum_{i=1}^n \sigma_i}{\sum_{i=1}^n G_i} \quad (11)$$

$$\sigma_{c_{dB}} = \frac{\sum_{i=1}^n G_i \sigma_{i_{dB}}}{\sum_{i=1}^n G_i} \quad (12)$$

III. Example

Let us array two antennas with gains of 61.7 and 56.1 dB and assume that both antennas are known to 0.1 dB one sigma. This approximates the DSN 64-meter and 34-meter antennas operating at S-band. Letting the subscript 1 denote the 64-meter antenna and 2 the 34-meter antenna, we get from Eq. (1)

$$G = G_1 + G_2 = 10^{(61.70/10)} + 10^{(56.10/10)} = 10^{(6.276)}$$

and from Eqs (3) and (7)

$$\begin{aligned} \sigma_u &= (\sigma_1^2 + \sigma_2^2)^{1/2} \\ &= \left\{ \left[\frac{\ln 10}{100} (10^{6.17}) \right]^2 + \left[\frac{\ln 10}{100} (10^{5.61}) \right]^2 \right\}^{1/2} \\ &= 35,326 \end{aligned}$$

Therefore, the gain of this two-antenna array is

$$G_{dB} = 62.76 \pm 0.081 \text{ dB (one sigma}_u\text{)}$$

or if we consider σ_c ,

$$\begin{aligned} \sigma_c &= \sigma_1 + \sigma_2 \\ &= \frac{\ln 10}{100} (10^{6.17}) + \frac{\ln 10}{100} (10^{5.61}) \\ &= 43,438 \end{aligned}$$

we get

$$G_{dB} = 62.76 \pm 0.100 \text{ dB (one sigma}_c\text{)}$$

We note that in this example, the gain of the array increased by 1.06 dB relative to the gain of the 64-meter antenna. However, even in the fully correlated error worst case, the uncertainty in the array gain has not increased over that of the individual antennas!

Figure 1 is a plot of the statistical (σ_u) and worst case (σ_c) gain error of this array with σ_1 held constant at 0.1 dB and σ_2 allowed to vary between 0 and 1 dB. From Fig. 1 we see that σ_2 can go to approximately 0.3 dB before the formal error of this two-antenna array reaches $\sigma_u = 0.1$ dB. From Fig. 1 we also see that even in the "worst case" the slope of the uncertainty is less than 1. That is,

$$\frac{d\sigma_c}{d\sigma_2} < 1$$

for

$$\sigma_1 = 0.1 \text{ dB}$$

This demonstrates that it is more critical to know the gain of the larger antenna of this two-element array. Indeed, it would be expected that the largest (highest gain) antenna in an array of unequal gain antennas would contribute most to the total array gain as well as contributing most to the uncertainty of the total array gain.

In Fig. 2, we see plotted the sum of the gains of a 64-meter (61.7-dB) antenna and zero to three 34-meter (56.1-dB) antennas. All antennas are assumed known to 0.1-dB one sigma. In this figure we see the uncorrelated error bars shrink as the number of antennas in the array is increased while the correlated error bars remain constant and the mean gain increases.

IV. Another Example

Let us look at the case of arraying a 64-m antenna (61.7 \pm 0.1 dB gain at S-band) with a second antenna whose gain will be allowed to vary downward from 61.7 dB, the uncertainty in the gain of the second antenna being kept constant at 0.1 dB. That is,

$$G_1 = 61.7 \pm 0.1 \text{ dB}$$

and

$$G_2 \leq 61.7 \pm 0.1 \text{ dB}$$

We now calculate that fraction of the uncertainty in the array gain due to the uncertainty in the gain of the smaller antenna. Looking at the uncorrelated case first, we get

$$U = \frac{\sigma_2}{\sigma_u} = \frac{\sigma_2}{(\sigma_1^2 + \sigma_2^2)^{1/2}} = \frac{(10^{G_{2dB}/10}) (\sigma_{2dB})}{\left\{ \left[(10^{G_{1dB}/10}) (\sigma_{1dB}) \right]^2 + \left[(10^{G_{2dB}/10}) (\sigma_{2dB}) \right]^2 \right\}^{1/2}}$$

Setting

$$G_{1dB} = 61.7$$

and

$$\sigma_{1dB} = \sigma_{2dB} = 0.1,$$

we get that

$$U = \frac{10^{G_{2dB}/10}}{\left[(10^{61.7})^2 + (10^{G_{2dB}/10})^2 \right]^{1/2}} \quad (13)$$

In the correlated case, we can write that

$$C = \frac{\sigma_2}{\sigma_c} = \frac{\sigma_2}{\sigma_1 + \sigma_2} = \frac{(10^{G_{2dB}/10}) (\sigma_{2dB})}{(10^{G_{1dB}/10}) (\sigma_{1dB}) + (10^{G_{2dB}/10}) (\sigma_{2dB})}$$

Again,

$$G_{1dB} = 61.7$$

and

$$\sigma_{1dB} = \sigma_{2dB} = 0.1$$

Then

$$C = \frac{10^{G_{2dB}/10}}{10^{6.17} + 10^{G_{2dB}/10}} \quad (14)$$

Equations (13) and (14) are shown plotted in Fig. 3 as a function of $G_1 - G_2$. In Fig. 3 we notice the apparently anomalous situation of U reaching 0.707 (instead of 0.5) when $G_1 = G_2$. This result is caused in the uncorrelated case because the

variances (σ_i^2) add while in the correlated case the standard deviations (σ_i) add.

V. Conclusion

From the preceding we see that (1) the array gain increases as antennas are added to the array, (2) the error bars do not grow under the constraint that we know the performance of all elements equally well, and (3) the knowledge of the 64-meter antenna performance is more important to the array than knowledge of the 34-meter antennas and therefore, we should spend most of our effort quantifying the larger antenna.

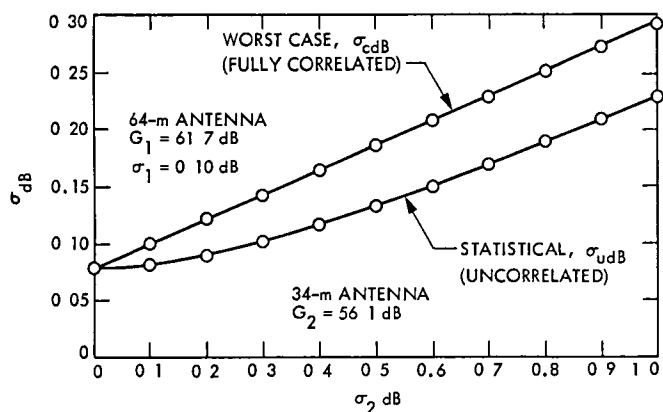


Fig. 1. Uncertainty in the array gain of a 64-meter antenna and a 34-meter antenna

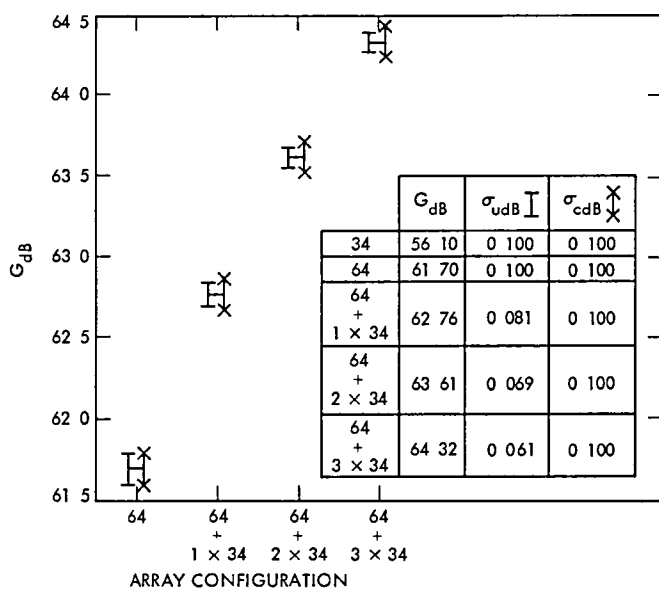


Fig. 2. Array gain vs configuration

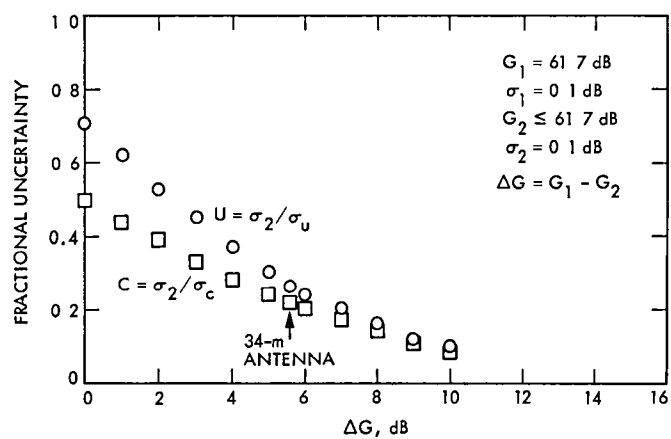


Fig 3 Fractional uncertainty in array gain due to smaller antenna

Life-Cycle Cost Analysis of Projects Using a Polynomial Cash Flow Model for Nonuniform Maintenance and Operations Costs

D S Remer

Communications Systems Research Section
and
Harvey Mudd College of Engineering Science

G Lorden

California Institute of Technology

A mathematical model is developed for calculating the life-cycle costs for a project where the maintenance and operations (M&O) costs change in a nonlinear manner with time. Closed-form solutions are presented for computing the present worth of projects with periodic cash flow profiles that can be approximated by polynomial functions. The results show that the life-cycle cost for a project can be grossly underestimated (or overestimated) if the M&O costs increase or decrease nonuniformly over time rather than being constant or linear as is often assumed in project economic evaluations. The following range of variables is examined (1) project life from 2 to 15 years, (2) interest rate from 0 to 30 percent per year, and (3) polynomials of order 0 to 5. Simplified solutions for the present worth are presented for two limiting cases: extended project lifetime and negligible interest rate. Also a simplified expression is provided for accurate present worth M&O estimates for DSN projects. In addition, a sensitivity analysis of the model based on graphical results and a numerical example plus tables and graphs are given to help the reader calculate M&O life-cycle costs over a wide range of variables.

I. Introduction

In the last few years, there has been an increasing emphasis on economic evaluations for comparing projects in the DSN. The total expected costs of systems or modifications, over their lifetimes, has become an important consideration to managers and engineers in TDA and the DSN. This interest has

recently culminated in the release of the Deep Space Network Life-Cycle Cost (LCC) Analysis Handbook and the Tracking and Data Acquisition Standard Practice for Life-Cycle Cost Analysis

The TDA Standard Practice states when an LCC analysis is required and what the analysis will contain, but does not state

how to do the analysis. The DSN LCC Analysis Handbook complements the Standard Practice by providing managers and engineers with the background information and guidance to conduct an LCC analysis. However, one of the major problems a user faces is modeling the maintenance and operations cost for the life of the system or systems modification. Usually one makes a zeroth-order assumption that the M&O costs are constant over time. This is an unrealistic assumption and has extremely limited use.

A first-order model was developed several years ago for M&O costs that vary in a linear manner with time (Refs. 1 and 2). The purpose of this paper is to develop closed-form solutions for the present worth of the maintenance and operations cost profile, based on an M^{th} -degree polynomial approximating the cash flows. For $M=0$ and $M=1$, this general model reduces to the usual models where one assumes that the M&O costs are constant over time ($M=0$) or where the M&O costs increase or decrease in a linear manner with time ($M=1$). Also, two limiting cases of special interest for quick engineering project estimates are presented. These are the cases of infinite project life and zero interest rate. The infinite life model approximates many extended life projects that are estimated to last well over 10 years. The zero-interest model may be used to approximate situations where the cost of capital is close to the inflation rate. The validity of this approximation was developed in Ref. 3 where it was shown that inflation and discounting largely cancel each other over long periods.

Application of the general polynomial model is discussed and results are presented graphically. A brief sensitivity analysis is made to show the relative importance of the costs of capital and the degree M of the cash flow polynomial. Finally, a simplified model for evaluating DSN projects is presented.

II. Previous LCC Mathematical Models

A. Zeroth-Order Model

The zeroth-order approximation that is often made in economic analysis is to assume that the M&O costs are constant each year. For this simple model, the present worth factor ($P/G_0, i, n$) of the life-cycle M&O costs for an n -year project and an interest rate i for a unit cash flow of \$1 per year is shown below and developed in Refs. 1 and 2

$$(P/G_0, i, n) = \frac{(1+i)^n - 1}{i(1+i)^n}, n > 1$$

The notation ($P/G_M, i, n$) will be used throughout the paper. P represents the present worth of the life-cycle costs.

G_M represents the order of the cash flow profile where $M=0$ is a constant cash flow, $M=1$ is a linear cash flow, etc. The i represents the interest rate used for the present worth calculation and n denotes the life of the project.

Needless to say the zeroth-order model has very limited application to most real problems where costs are rarely constant over time. In the next section we will consider a first-order model.

B. First Order Model

A mathematical model for calculating the life-cycle costs of a project where the M&O costs increase or decrease in a linear manner with time has been treated in detail in Refs. 1 and 2.

From Refs. 1 and 2, the present worth factor for the life-cycle M&O costs with a linear increasing cash flow profile for an n -year project and an interest rate i was shown to be

$$(P/G_1, i, n) = \frac{(1+i)^n - ni - 1}{i^2(1+i)^n}, n > 1$$

The results in Refs. 1 and 2 show that the life-cycle cost for a project can be grossly underestimated (or overestimated) if the operating costs increase (or decrease) uniformly over time rather than being constant as is often assumed in project economic evaluations. This model is a good first step forward for analyzing life-cycle M&O costs, but a more general model is really needed to supplement Appendix E in the DSN Life-Cycle Cost Analysis Handbook. This appendix is entitled, "Tools for Calculating a Life-Cycle Cost." The following theoretical development will substantially improve the tools available to do a life-cycle cost analysis for most cash flow profiles that we will probably encounter in the DSN. It will no longer be necessary to make such simplifying assumptions as uniform annual M&O costs or linear M&O costs. The following model, which is based on the recent developments in Ref. 4, will allow us to calculate the life-cycle M&O costs for a generalized cash flow profile.

III. A General Life-Cycle Cost Mathematical Model

A. Formulation of the Model

The cash flow profile for an n -year project can often be approximated by some M^{th} degree polynomial of the form

$$y = C_M(x-1)^M + C_{M-1}(x-1)^{M-1} + \dots + C_1(x-1) + C_0 \quad (1)$$

where C_0, C_1, \dots, C_M are constants and y is the amount of cash flow in year x . We assume that this profile can be represented as a series of discrete annual payments (see Fig 1). Throughout this paper, discrete compounding of interest is used and, for simplicity, cash flows are assumed to be in dollars. We use $(x-1)$ instead of x in Eq (1) to be consistent with standard engineering economic texts (Refs 5 and 6) where the cash flows for a linear gradient have their first costs occurring at the end of year 2.

All results are based on annual cash flows and an annual interest rate. However, the results may be extended to any type of periodic cash flow if an appropriate interest rate is used. For example, the above polynomial may be used to describe cash flows on a monthly basis, where n represents the number of months in the project lifetime. The subsequent analysis would then require a monthly interest rate.

We will first consider the most basic form of Eq (1). In this case, the cash flow y in year x of a specific n -year project is determined by the equation

$$y = (x-1)^M \quad 1 < x < n \quad (2)$$

for a suitable choice of integer $M \geq 1$ (refer to Fig 2). The present worth in year zero of the total cash flow of this project at an annual interest rate i is

$$(P/G_M, i, n) = \sum_{x=2}^n (x-1)^M (1+i)^{-x} \quad (M \geq 1) \quad (3)$$

We will now construct a more tractable equation for the present worth of the M&O life-cycle costs. Equation (3) reduces to the following after algebraic manipulation

$$(P/G_M, i, n) = \frac{1}{i} \sum_{x=1}^n [\nabla(x^M)] (1+i)^{-x} \frac{n^M}{i(1+i)^n} \quad (M \geq 1) \quad (4)$$

where ∇ is the backward difference operator

$$\nabla(x^M) = x^M - (x-1)^M = \sum_{k=1}^M \binom{M}{k} x^{M-k} (-1)^{k+1} \quad (5)$$

and $\binom{M}{k}$ is the binomial coefficient

$$\frac{M!}{k!(M-k)!}$$

By substituting Eq. (5) into Eq (4), we have

$$(P/G_M, i, n) = \frac{1}{i} \sum_{x=1}^n \sum_{k=1}^M \binom{M}{k} x^{M-k} (-1)^{k+1} (1+i)^{-x} - \frac{n^M}{i(1+i)^n}$$

which reduces to the recursive formula

$$(P/G_M, i, n) = \frac{1+i}{i} \sum_{k=0}^{M-1} \binom{M}{k} (-1)^{M-k-1} (P/G_k, i, n) + \frac{(-1)^M}{i} - \frac{(n-1)^M}{i(1+i)^n}, \quad M \geq 1 \quad (6)$$

Thus, $(P/G_M, i, n)$ is a linear combination $(P/G_0, i, n)$, $(P/G_1, i, n)$, \dots , $(P/G_{M-1}, i, n)$.

In Table 1, we present closed-form solutions for $(P/G_M, i, n)$, when $M = 0, 1, 2, 3, 4$, or 5.

We feel that using $M \leq 3$ will provide sufficient accuracy for most LCC project evaluations in the DSN. In addition, this model will be useful for doing sensitivity and risk analyses for LCC studies.

B. Special Cases

There are two cases of special engineering interest. The first concerns the limiting behavior of the present worth as the project life n becomes infinite. This behavior is important when evaluating projects that have long lifetimes of more than, say, 20 years, like a DSN antenna. The second case is when the interest rate approaches zero. This can be used for doing quick engineering calculations where the interest and inflation rates nearly cancel. This was shown to be the case for the DSN (Ref. 3).

1. Infinite project life. The effect of discounting causes the present worth $(P/G_M, i, n)$ of a project to converge as the lifetime n becomes infinite. This can be verified by applying the ratio test to the terms of the infinite series generated from Eq (3). Thus, $(P/G_M, i, n)$ exhibits asymptotic behavior as

$n \rightarrow \infty$ If the interest rate is not zero, a recursive formula for this asymptotic level is

$$F(M, i) = \lim_{n \rightarrow \infty} (P/G_M, i, n) = \frac{1+i}{i} \sum_{j=1}^{M-1} \binom{M}{j} (-1)^{M-j-1} F(j, i) + \frac{(-1)^M}{1}, \quad M \geq 1$$

where $F(0, i) = 1/i$. This formula was derived by taking limits in Eq. (6). In Table 2, the limiting values are summarized for $M = 0, 1, 2, 3, 4$, and 5.

These results are very useful for evaluating facilities that are expected to have a long lifetime, like DSN antennas.

2. Zero interest rate. The second case of special interest in the DSN concerns the limiting behavior of the present worth as the interest rate i approaches zero. We will now develop closed-form solutions for these limiting cases. We can avoid taking limits of the closed-form expressions in Table 1 by considering the simpler case of Eq. (3). Using this equation, the present worth of the zero-interest form of the basic model for $M \geq 1$ is

$$(P/G_n, 0\%, n) = \sum_{x=2}^n (x-1)^M = \sum_{x=1}^{n-1} x^M$$

In this special case, the present worth is just the sum of the M^{th} powers of the first $n-1$ positive integers; closed-form expressions for this sum are widely available (Ref. 7). Based on our definition of $(P/G_0, i, n)$, this present worth is n when $M=0$. In Table 2, the limiting values are summarized for $M = 0, 1, 2, 3, 4$, and 5.

C. Application of the LCC Model

We will now discuss how one applies the model and then give an example. As described in subsection III-A, the predicted cash flow profile of an n -year project can be approximated, in general, by some M^{th} -degree polynomial of the form given in Eq. (1). Consequently, the cash flow y in year x has $M+1$ components; a constant cost of size C_0 , a linear cost of size C_1, \dots , and a M^{th} -degree cost of size C_M . Hence, we can determine the total present worth PW_{LCC} of a project having this cash flow profile by finding the corresponding linear combination of basic present worths $(P/G_0, i, n)$, $(P/G_1, i, n), \dots, (P/G_M, i, n)$

$$PW_{LCC} = C_M(P/G_M, i, n) + \dots + C_1(P/G_1, i, n) + C_0(P/G_0, i, n)$$

This method will be described in more detail in subsection III-D.

In most real-life situations, the future cash flow profile is determined by an equation of the form

$$y = B_M x^M + B_{M-1} x^{M-1} + \dots + B_1 x + B_0 \quad (7)$$

where y is the cash flow in year x . Accordingly, Eq. (7) is converted to the required form of Eq. (1) by using the identity

$$C_j = \sum_{k=j}^M B_k \binom{k}{j}, \quad j = 0, 1, 2, \dots, M$$

It should be noted that repeated applications of this model may be necessary to evaluate the present worth of the total LCC. For example, suppose we found polynomials y_1 and y_2 corresponding to the cash flow profiles for initial investment costs and for annual maintenance and operations costs, respectively, as shown in Fig. 3. Since the initial investment costs begin in year 1, we may directly apply the general model to the cash flow polynomial y_1 to get R_1 , the present worth in year zero of the initial investment costs. Next, we apply the general model to y_2 to get the present worth in year n' of the annual operating costs. We then apply the present worth factor $(P/F, i, n') = 1/(1+i)^{n'}$ to the result to get R_2 , the present worth in year zero of the annual operating costs. Finally, the present worth in year zero of the total LCC is the sum of R_1 and R_2 .

The polynomial form that one uses to represent the yearly cash flows can be developed either by curve-fitting or from a forecasting model.

Equivalent uniform annual cost (EUAC) computations can easily be performed by converting the previous present worth calculations to a level annuity:

$$EUAC = (P/G_M, i, n) (P/G_0, i, n)^{-1}$$

where

$$(P/G_0, i, n)^{-1} = i(1+i)^n / [(1+i)^n - 1]$$

is the capital recovery factor at interest rate i .

D. Example

This section gives an example of how one might apply the above model in an economic analysis of a proposed project in the DSN. Let's consider, as an illustration, the introduction of

increased automation at a Deep Space Station, to reduce operating costs (Refs 8 through 11). A strong effort has been made since 1969 to reduce operating costs by reducing the number of people at the station. For example, the crew size at a station has been reduced from 26 people in 1969 to 4 today. However, this automation requires large capital investments that must be justified using an LCC analysis. The capital expenditures required must be compared to the reduction in future operating and maintenance costs. Our following example will look at the present value of the projected M&O cash flows for an existing system during its entire life.

Suppose, for example, that the future operating costs of one part of the Deep Space Station are expected to vary according to the second-degree model

$$y = 6x^2 - 11x + 10$$

where y is the operating costs (in units of \$10,000 for this example) at year x , and year zero is taken to be the startup time of the automation project. This polynomial has the form of Eq (7) and thus is rewritten as

$$y = 6(x - 1)^2 + (x - 1) + 5$$

to coincide with the form of Eq. (1). The present worth of the operating costs at 10 years after startup is

$$\begin{aligned} PW_{LCC} &= 6(P/G_2, 10\%, 10) + (P/G_1, 10\%, 10) \\ &\quad + 5(P/G_0, 10\%, 10) \end{aligned}$$

Using the formulas in Table 1 with $i = 0.10$ and $n = 10$, we have

$$(P/G_0, 10\%, 10) = 6.1446$$

$$(P/G_1, 10\%, 10) = 22.8913$$

$$(P/G_2, 10\%, 10) = 133.7292$$

Therefore,

$$\begin{aligned} PW_{LCC} &= 6(133.7292) + 22.8913 + 5(6.1446) \\ &= 8.56 \text{ million dollars} \end{aligned}$$

Therefore, if our implementation cost for automation is less than 8.56 million dollars, the expenditure is justified.

Compared to conventional discounting of individual cash flows, this model is very helpful when different polynomials

for cash flow prediction exist in several subsystems of one complex system. For example, in a Deep Space Station there will be different cash flow polynomials for the antenna, receivers, and transmitters.

IV. Graphical Review of Results

We examined the behavior of the general present-worth $(P/G_M, i, n)$ as a function of three parameters: the order M of the cash flow model, the prevailing cost of capital i , and the project life n .

First, using the formulas in Table 1, $(P/G_M, 10\%, n)$ was calculated for values of M between 0 and 5 and for project lifetimes of from 2 to 15 years. These results are shown in Fig 4. For $M = 0$ and 1, the graphs will quickly approach the asymptotic limits described in Table 2.

Next, we performed a graphical sensitivity analysis of the relative effects of the parameters i and M on the present-worth function $(P/G_M, i, n)$, $M = 0, 1, 2, 3$. For each of these values of M , semilogarithmic plots similar to those in Fig 4 were made for interest rates of 5%, 10%, and 15%. These results are presented in Fig 5 and show four families ($M = 0, 1, 2, 3$) of present-worth curves at three interest rates (5%, 10%, 15%). We can see from Fig 5 that M has a more significant impact on the LCC present value $(P/G_M, i, n)$ than does the interest rate i . For example, increasing M from 2 to 3 increases the present worth by a factor of about 6, for interest rates of 5% or 10%. However, decreasing i from 10% to 5% increases the present worth only by a factor of about 1.4, for $M = 2$, or 3.

The next section deals with approximating the logarithm of these curves by linear functions, for a 10-year project life. The approach that will be presented is valid for any project life, but a 10-year project life is of interest to the DSN and several U.S. government agencies, where it is used as a benchmark in evaluating many projects.

V. Simplified Model for a Project With a Ten-Year Life Cycle

Figure 6 is a semilogarithmic graph of $(P/G_M, i, 10)$ as a function of the interest rate i (1% to 15%). Since the six plots in this figure are nearly straight lines, $\log_{10} (P/G_M, i, 10)$ may be approximated by a linear function for each value of M . Consequently, the present worth may be approximated by much simpler formulas than those given in Table 1. Six linear regressions on $\log_{10} (P/G_M, i, 10)$ were performed and the resulting functions $f_M(i)$ that approximate $(P/G_M, i, 10)$ are presented in Table 3. Clearly, $\log_{10} f_M(i)$ produces a straight line for each M .

The functions in Table 3 are very good approximations of the complex formulas in Table 1. The maximum relative error is only 1.7% and the average relative error is only 0.7%. Returning to the example in subsection III-F, we have

$$PW_{LCC} = 6(P/G_2, 10\%, 10) + (P/G_1, 10\%, 10) \\ + 5(P/G_0, 10\%, 10)$$

We now use the values given in Table 3 to get

$$PW_{LCC} = 6f_2(10\%) + f_1(10\%) + 5f_0(10\%) \\ = 6S_2T_2^{0.1} + S_1T_1^{0.1} + 5S_0T_0^{0.1} \\ = 8.62 \text{ million dollars}$$

(vs. 8.56 million dollars as calculated in subsection III-F using the exact equation). The relative error between these figures is only 0.7%. This is well within the accuracy of most project economic models.

Most U.S. government agencies, such as NASA, the Navy, and the Department of Defense, use a value of $i = 10\%$ in their calculations (Ref. 12). Table 4 gives the value of $f_M(i)$ for $i = 10\%$. Now it is very easy to estimate the life-cycle cost for projects in the DSN by simply using the results in Table 4.

VI. Summary

We have constructed a mathematical model for determining the present worth of the life-cycle costs of a project when the expected future disbursements can be approximated by a general polynomial function, and we have illustrated by means of an example the application of the model. This model is a generalization of a basic model that has the future cash flows predicted by the equation $y = (x - 1)^M$, where y is the expected costs of the proposed project of year x . The model extends the usual engineering economic models of a level annuity ($M = 0$) and a linear gradient ($M = 1$). We then developed closed-form expressions of $(P/G_M, i, n)$ for $M = 0$ through $M = 5$.

Limiting cases of the model were also examined, and we found that in the case of an infinite project life or a negligible interest rate, the present worth calculation could be considerably simplified. A method for computing equivalent uniform annual costs was also presented.

The present worth $(P/G_M, i, n)$ is a function of the interest rate i , the estimated project life n , and the degree M of the predicting equation. A graphical sensitivity analysis of the model showed that the parameter M corresponding to the degree of the predicting equation was more influential in determining the present worth $(P/G_M, i, n)$ than the interest rate i . For a project with a 10-year life, graphical analysis showed that approximations of the present worth expressions could be developed using linear regression. These approximations are much simpler than the original expressions and have the advantage of introducing an average error of only 0.7%.

By using the results developed here for the DSN, the life-cycle cost calculation will be greatly simplified when comparing projects.

References

- 1 Remer, D S., "A Life-Cycle Cost Economics Model for Projects with Uniformly Varying Operating Costs," in *The Deep Space Network Progress Report 42-39*, pp. 60-70, Jet Propulsion Laboratory, Pasadena, Calif , June 15, 1977.
- 2 Remer, D S , "A Life-Cycle Cost Economics Model for Automation Projects with Uniformly Varying Operating Costs," *National Telecommunications Conference Record*, 32 4-1 to 32 4-6, 1977
- 3 Eisenberger, I , Remer, D S , and Lorden, G , "The Role of Interest Rates and Inflation Rates in Life-Cycle Cost Analysis," in *The Deep Space Network Progress Report 105-109*, pp 60-70, Jet Propulsion Laboratory, Pasadena, Calif , February 15, 1978
- 4 Almond, B , and Remer, D S , "Present Worth Analysis of Capital Projects Using a Polynomial Cash Flow Model," *Engineering Costs and Production Economics*, 5, pp 33-42, 1980
- 5 Grant, E L , Ireson, W G , and Leaveworth, R S , *Principles of Engineering Economy*, 6th ed , New York The Ronald Press Co , 1976
- 6 Tarquin, A J., and Blank, L T , *Engineering Economy*, New York McGraw-Hill, Inc , 1976
- 7 Selby, S M , ed , *Standard Mathematical Tables*, 23rd ed , Cleveland CRC Press, Inc , 1975
- 8 Remer, D S , Eisenberger, I , and Lorden, G , "Economic Evaluation of DSS 13 Unattended Operations Demonstration," in *The Deep Space Network Progress Report 42-45*, pp 165-171, Jet Propulsion Laboratory, Pasadena, Calif , March and April 1978
- 9 Remer, D S , and Lorden, G , "Initial Economic and Operations Data Base for DSS 13 Automation Test," in *The Deep Space Network Progress Report 42-49*, pp 78-85, Jet Propulsion Laboratory, Pasadena, Calif , November and December 1978
- 10 Remer, D S , and Lorden, G , "Preliminary Maintenance Experience for DSS 13 Unattended Operations Demonstration," in *The Deep Space Network Progress Report 42-51*, pp 150-155, Jet Propulsion Laboratory, Pasadena, Calif , March and April 1979
- 11 Remer, D S , and Lorden, G , "Deep Space Station Automation Demonstration," in *The Deep Space Network Progress Report 42-57*, pp 103-119, Jet Propulsion Laboratory, Pasadena, Calif , March and April 1980
- 12 *Economic Analysis Handbook*, Department of the Navy, Naval Facilities Engineering Command, NAVFAC, p 442, June 1975

Table 1. Closed-form solutions for the basic model ($n > 1$)

M	$(P/G_M, i, n)$
0	$\frac{1}{i(1+i)^n} [(1+i)^{n-1} - 1]$
1	$\frac{1}{i^2(1+i)^n} [(1+i)^n - ni - 1]$
2	$\frac{1}{i^3(1+i)^n} [(1+i)^{n+1} + (1+i)^n - n^2i^2 - (2n+1)i - 2]$
3	$\frac{1}{i^4(1+i)^n} [(1+i)^{n+2} + 4(1+i)^{n+1} + (1+i)^n - n^3i^3 - (3n^2 + 3n + 1)i^2 - 6(n+1)i - 6]$
4	$\frac{1}{i^5(1+i)^n} [(1+i)^{n+3} + 11(1+i)^{n+2} + 11(1+i)^{n+1} + (1+i)^n - n^4i^4 - (4n^3 + 6n^2 + 4n + 1)i^3 - (12n^2 + 24n + 14)i^2 - (24n + 36)i - 24]$
5	$\frac{1}{i^6(1+i)^n} [(1+i)^{n+4} + 26(1+i)^{n+3} + 66(1+i)^{n+2} + 26(1+i)^{n+1} + (1+i)^n - n^5i^5 - (5n^4 + 10n^3 + 10n^2 + 5n + 1)i^4 - (20n^3 + 60n^2 + 70n + 30)i^3 - (60n^2 + 180n + 150)i^2 - (120n + 240)i - 120]$

Table 2 Asymptotic levels for the basic model

M	Infinite project life, $\lim_{n \rightarrow \infty} (P/G_M, i, n)$	Zero interest rate, $\lim_{i \rightarrow 0} (P/G_M, i, n)$
0	$1/i(1+i)$	$n - 1$
1	$1/i^2$	$(n^2 - n)/2$
2	$(i + 2)/i^3$	$(2n^3 - 3n^2 + n)/6$
3	$(i^2 + 6i + 6)/i^4$	$(n^4 - 2n^3 + n^2)/4$
4	$(i^3 + 14i^2 + 36i + 24)/i^5$	$(6n^5 - 15n^4 + 10n^3 - n)/30$
5	$(i^4 + 30i^3 + 150i^2 + 240i + 120)/i^6$	$(2n^6 - 6n^5 + 5n^4 - n^2)/12$

Table 3 Approximating functions $f_M(l)$ for $(P/G_M, l, 10)$

M	$(P/G_M, l, 10) \approx f_M(l) = S_M \cdot T_M^l$	
	$S_M/10^M$	$T_M \times 10^{-5}$
0	9 739	108
1	4 390	16 01
2	2 785	7 008
3	1 981	4 228
4	1 501	3 033
5	1 183	2 403

Table 4 Approximations for $(P/G_M, 10\%, 10)$

M	$(P/G_M, 10\%, 10) \approx f_M(10\%)$	
	$f_M(10\%)/10^M$	Relative error, %
0	6 193	0 78
1	2 306	0 75
2	1 347	0 71
3	0 911	0 68
4	0 608	0 66
5	0 514	0 65

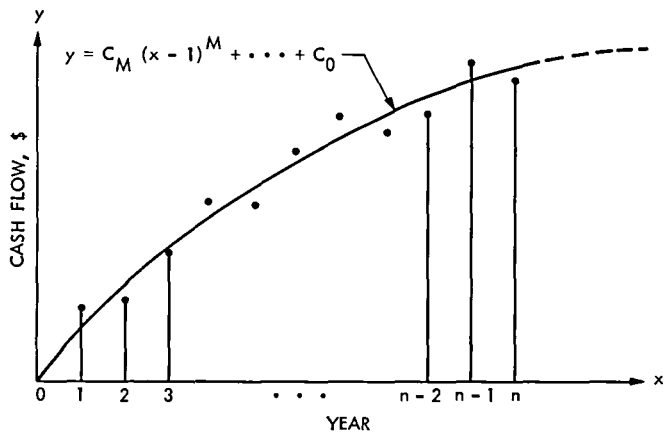


Fig 1. General approximation of cash flows

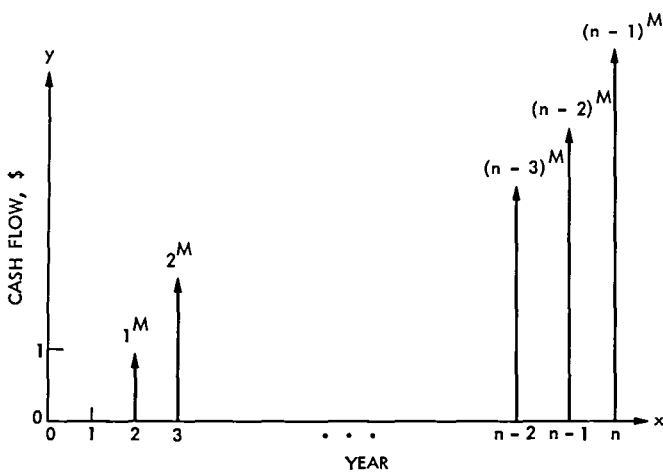


Fig 2. Cash flow profile for the basic model

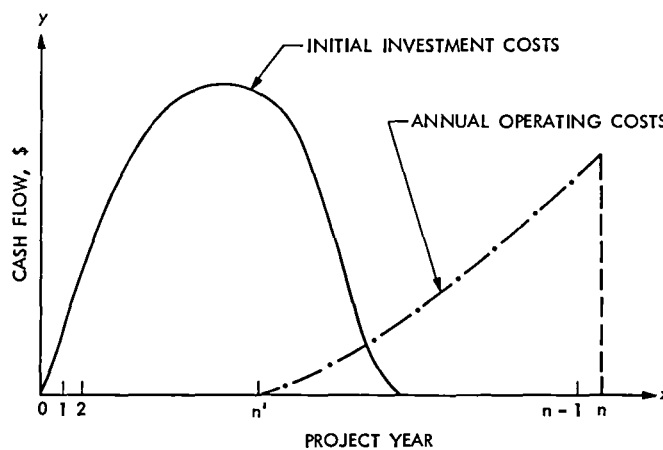


Fig 3. Two-stage life-cycle cost example

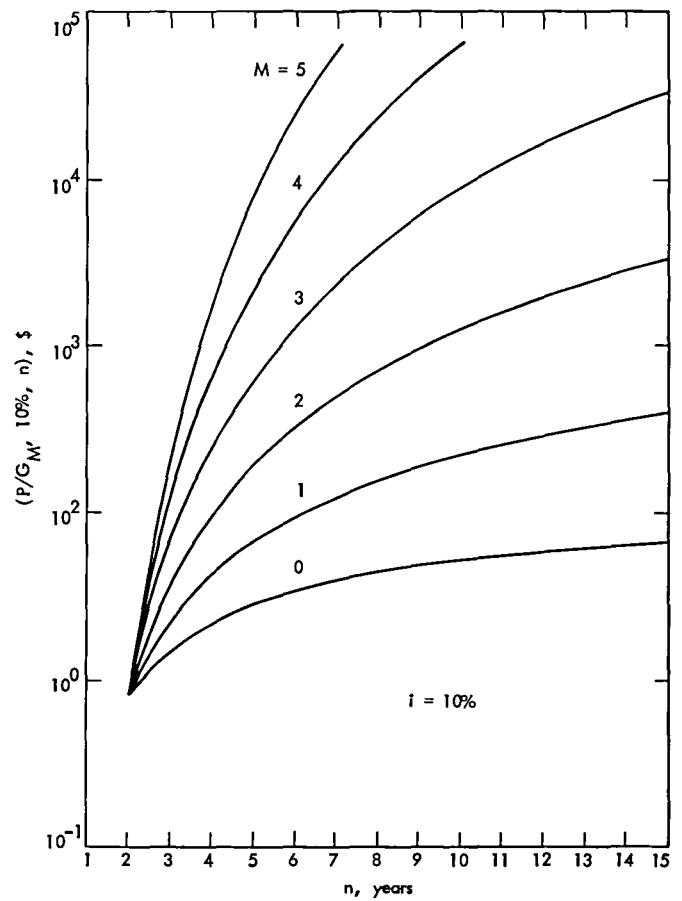


Fig 4. Life-cycle cost present worth $(P/G_M, 10\%, n)$ vs project life

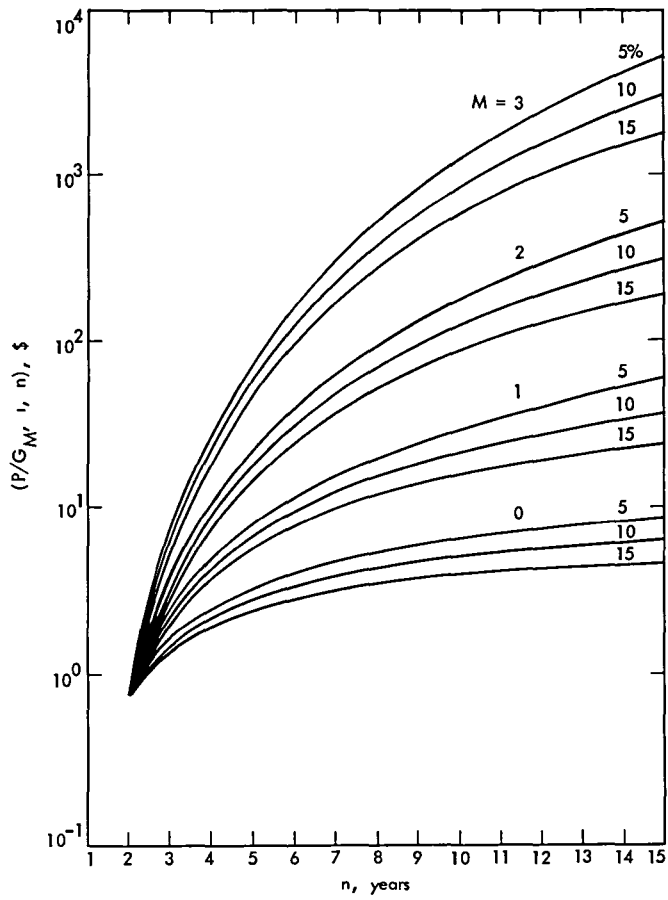


Fig. 5. Life-cycle cost present worth $(P/G_M, i, n)$ vs project life

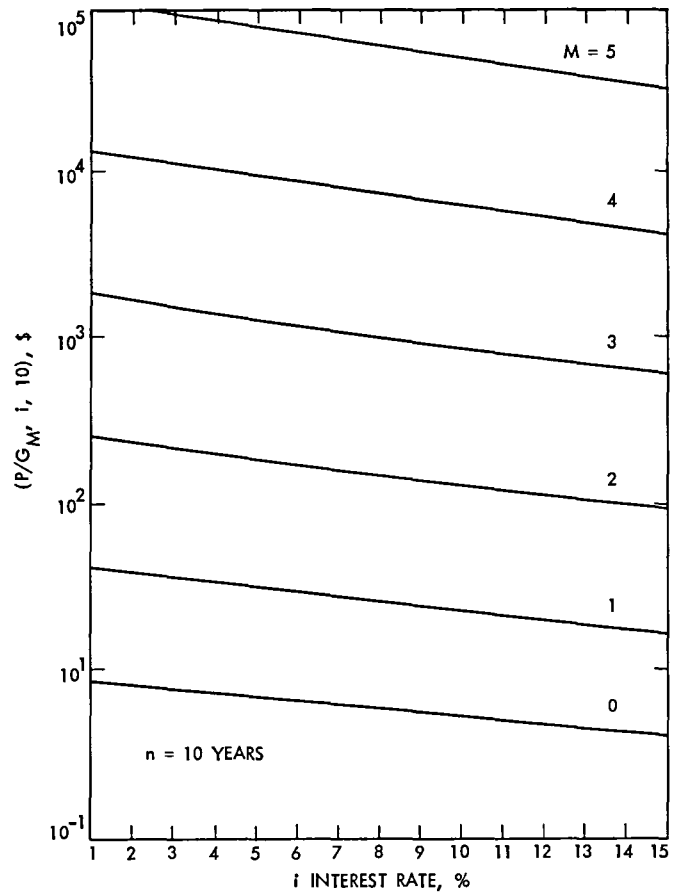


Fig. 6. Life-cycle cost present worth vs interest rate

Sine Wave Ranging Revisited

A. I. Zygielbaum
DSN Data Systems Section

A. Robinson
DSS 43, Tidbinbilla, Australia

It has been conjectured that DSN ranging accuracy would be improved if the range code were a sine wave. In this article, measurements are presented which demonstrate that this is not the case but that the use of sine waves may be worthwhile to conserve the uplink frequency spectrum.

I. Introduction

DSN ranging precision was improved by a factor of 4 by using a 1-MHz initial range code versus the 500-kHz code used earlier (Refs. 1 and 2). The range codes are square waves modulated onto a carrier transmitted to and returned from a spacecraft via a transponder. Comparison of the phase of the transmitted wave to that returned yields a direct measure of time-delay and an indirect measure of range. Square wave ranging suffers from waveform distortion due to asymmetric amplitude and phase distortion in the communications channel. The problem is compounded by a mismatch between the actual correlation function produced by the ranging hardware and that assumed by the software. In essence, the software assumes that the returned signal is a square wave while limitation of the DSN transmitter, spacecraft transponder, and, to a lesser extent, the DSN receiver, reduces the signal to a badly distorted sine wave.

As explained by Layland, et al., in Ref. 1, the ± 1.5 -MHz transponder bandwidth passes the 500-kHz code and a distorted version of its third harmonic. The resulting signal is at best a poor sine wave. The same transponder will, however, propagate only the fundamental of 1-MHz code. The resulting

signal is received by the DSN and filtered again to eliminate regenerated harmonics. Hence a simple sine wave is applied to the ranging system which uses a sine wave correlation model. Therefore, there is little waveform distortion or mismatch in correlation.

The fundamental concept is that the most accurate ranging comes from a signal with the simplest spectrum. It was conjectured therefore that the uplink carrier should be modulated with a sine wave instead of a square wave, thereby eliminating the need for filters. This article reports an analytical and empirical study of that conjecture.

II. An Analysis of the Modulated Signal Spectrum

A. Square Wave Modulation Spectrum

A carrier phase modulated by a square wave may be represented by the equation

$$S_T(t) = A \sin [\omega_c t + \phi_c + k \text{SIN}(\omega_m t)] \quad (1)$$

where

A is amplitude

ω_c is the carrier frequency

ϕ_c is an arbitrary carrier phase

k is the modulation amplitude or "index" and is defined as the peak phase excursion

ω_m is the modulation frequency

and

SIN is the square wave approximation of the sine function

Equation (1) may be simplified by setting $A = 1$, $\phi_c = 0$ and by some trigonometric manipulation to

$$S_T(t) = \cos(k) \sin(\omega_c t) + \cos(\omega_c t) \sin(k) \text{SIN}(\omega_m t) \quad (2)$$

where we have used the relation

$$\sin[k \text{SIN}(\omega_m t)] = \sin(k) \text{SIN}(\omega_m t)$$

A Fourier transform of (2) can easily be made since the DSN ranging code is, for ease of mechanization, an integer submultiple of the carrier frequency. This fortuitously allows representation by a discrete series.

Because of the mechanization, we may write

$$\omega_c = 2\pi f_c = \frac{2\pi\alpha}{T_0}$$

where T_0 is the period of the modulation and α is the ratio of the modulation frequency to the carrier frequency.

Given the form

$$S_T(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos \frac{2\pi n t}{T_0} + \sum_{n=1}^{\infty} B_n \sin \frac{2\pi n t}{T_0}$$

$$A_0 = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} S_T(t) dt$$

$$A_n = \frac{2}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} S_T(t) \cos \frac{2\pi n t}{T_0} dt$$

$$B_n = \frac{2}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} S_T(t) \sin \frac{2\pi n t}{T_0} dt$$

for the Fourier series and with the knowledge that α is an integer, we can easily, but tediously, derive that for $\alpha \neq n$

$$A_0 = 0$$

$$A_n = 0$$

$$B_n = \frac{1}{\pi} \sin k \left[\frac{1 - \cos \pi(n - \alpha)}{n - \alpha} \right]$$

hence the "power" in the component is

$$P_n = A_n^2 + B_n^2 = B_n^2.$$

For

$$\alpha = n$$

$$A_0 = 0$$

$$A_n = 0$$

$$B_n = \cos k$$

the power is $P_n = \cos^2 k$

Clearly, the case $\alpha = n$ is the carrier term, $(n - \alpha) = \pm 1$ is the fundamental code sideband, $(n - \alpha) = \pm 2$ is the sideband of the second harmonic of the code, etc. Note that for $(n - \alpha)$ even, $B = 0$. Hence only odd harmonics are present. Therefore, the carrier amplitude is

$$C = \cos k,$$

the fundamental frequency ($n - \alpha = 1$) amplitude is

$$C_1 = \frac{1}{\pi} \sin k \left[\frac{1 - \cos \pi}{1} \right] = \frac{2}{\pi} \sin k$$

and so on

Standard practice is to give the modulation index as a carrier suppression in decibels of power. Given that the unmodulated carrier has amplitude 1, the carrier suppression is simply

$$\delta = 20 \log_{10} \cos k$$

Table 1 summarizes the relative power of the carrier and sidebands to total spectral power (which is arbitrarily set to 1)

B. Sine Wave Modulation Spectrum

Many textbooks give this elementary analysis. From the book by Taub and Schilling (Ref 3), one may write

$$S_T(t) = \cos [\omega_c t + k \sin (\omega_m t)]$$

and the Fourier series representation is simply

$$\begin{aligned} S_T(t) = & J_0(k) \cos \omega_c t \\ & - J_1(k) [\cos (\omega_c - \omega_m)t - \cos (\omega_c + \omega_m)t] \\ & + J_2(k) [\cos (\omega_c - 2\omega_m)t + \cos (\omega_c + 2\omega_m)t] \\ & - J_3(k) [\cos (\omega_c - 3\omega_m)t - \cos (\omega_c + 3\omega_m)t] + \dots \end{aligned}$$

where $J(k)$ is the Bessel function of the first kind of order n and the power of each component is in arbitrary units

$$P_0 = (J_0(k))^2$$

$$P_1 = (J_1(k))^2$$

$$P_2 = (J_2(k))^2$$

Carrier suppression is given by $\delta = 20 \log_{10}(J_0(k))$ for the total power arbitrarily set to 1. Carrier suppression was converted to modulation index or peak phase excursion by a piecewise approximation to the inverse Bessel function. A Texas Instruments TI-59 calculator program using a Newtonian iterative approximation provided the conversion

Table 2 gives the relative power of each sideband for various carrier suppressions

III. Comment on Analysis of Range Error

As reported by Layland et al, in Refs 1 and 4 and in an analysis by the author, asymmetric phase delay and amplitude across the ranging channel can lead to significant errors in the range measurement. Depending on the response of the channel, errors on the order of several tens of nanoseconds are possible. Rather than reiterate the analysis, actual range error measurements will be reported below. A complete analysis would require extensive and complex measurement of the communications equipment and is beyond the scope or need of this presentation.

IV. Test Configuration

In order to test the conjecture that transmitting a sine wave range code will eliminate range error, the MU2 R&D ranging system was used at DSS 43 to obtain station delay measurements. The basic test configuration appears in Fig 1. Range code generated by the range system is applied to the carrier by the exciter modulator. After amplification by the transmitter, the signal is echoed to the receiver by the test translator, which is a wide bandwidth device in comparison to the range channel. The receiver provides the returned signal to the MU2.

During the tests discussed in this article, the 1-MHz ranging code, a hard square wave, was applied directly or through filter f_1 to the Block IV exciter. When applied directly, the uplink signal was in fact modulated by a squarewave. Filter f_1 , a low-pass filter with cutoff at about 1.5 MHz, scrubbed all but the fundamental from the signal when sine wave modulation was required. Figure 2 presents the response of filter f_1 .

The received range code is presented to the MU2 as a phase-modulated signal riding on a 10-MHz IF carrier. As mentioned earlier, a 3-MHz passband filter, f_2 , may also be used to scrub the code harmonics so that a pure sine wave range code is correlated by the MU2.

The MU2 was used in a mode where range delay measurements were made continuously at discrete intervals, while the local correlation model or reference code was stepped in phase. The resulting data portrays the correlation function of the correlator. The phase step size is precisely known. An ideal system would show that same phase difference between successive range delay measurements. Any deviations would be due to distortion in the system.

Spectral measurements were made with a Hewlett-Packard HP851A/8551A spectrum analyzer. As shown in Fig 3, the transmitted signal was sampled immediately prior to radiation from the antenna horn

V. Range Accuracy Results

Four test cases are presented in Figs 4(a) through (d). These show 1/16 of the 0 to 2π range of possible phase differences between the incoming code and local code model. The abscissa is marked in angles of phase difference. The ordinate gives the difference between the actual phase difference and that measured by the MU2 in units of time (1 μ sec is approximately 2π radians for the 1-MHz code). One sixteenth of the total possible range is sufficient because the same pattern repeats due to symmetries in the correlation function.

(1) Case 1 No Filters (Fig 4a)

By using neither filter f_1 nor f_2 , the system is transmitting and receiving square waves. The square wave correlation results in about 9.5 ns of peak error.

(2) Case 2 Uplink Filter Only (Fig 4b)

In this case, a sine wave is assumed for the correlation model. The uplink filter assures that the applied modulation is a sine wave. Note that the peak error is about 5.5 ns. Clearly waveform distortion still exists.

(3) Case 3 Receive Filter Only (Fig 4c)

The correlation assumes a sine wave in this case because the receive filter propagates only the fundamental frequency of the range code. Note that the peak error is only 1 ns. In fact, this is the MU2 quantization error.

(4) Case 4 Uplink and Receive Filters (Fig 4d)

The error in this case is also about 1 ns, proving that at least the uplink filter causes no harm.

VI. Results of Spectral Measurements

Figure 5 shows the spectrum of the uplink signal with no uplink filter at various carrier suppressions. This spectrum should represent the coefficients in Table 1. One can immediately note that even harmonic sidebands appear. This is due to the noninfinite bandwidth and, to a lesser degree, to distortion.

Figure 6 displays the spectrum of the uplink given the use of the uplink filter. Hence the modulating range code is a sine wave. Comparison with Table 2 will reveal small deviations from the ideal. The more subtle of these may be errors in the spectrum analyzer. Nevertheless, amplitude asymmetries are apparent.

VII. Conclusions

Several conclusions may be drawn from the results. First it is apparent that nonlinearities in the modulator, exciter, transmitter, and/or receiver distort the ranging channel. Hence sine wave modulation yields virtually no improvement in ranging accuracy. Clearly, however, filtering the downlink affords a significant reduction in distortion effects. But at least uplink filtering does not degrade performance and comparison of Figs. 5 and 6 does show that the uplink spectrum can be conserved by uplink filtering. Conservation will become increasingly important as the range code frequency is increased to achieve greater range precision.

References

1. Layland, J. W., Zygielbaum, A. I., and Hubbard, W. P., "On Improved Ranging," *DSN Progress Report 42-46*, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1978.
2. Zygielbaum, A. I., "Installation of the MU2 Ranging System in Australia," *DSN Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1979.
3. Taub, H., and Schilling, D. L., *Principles of Communication Systems*, McGraw-Hill, Inc., New York, 1971.
4. Martin, W. L., and Layland, J. W., "Binary Sequential Ranging with Sine Waves," *DSN Progress Report 42-31*, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1976.

Table 1. Sideband relative to total power for square wave modulation (Note: $P \sim 0$ for n odd)

Carrier suppression, dB	k	P_0	P_1	P_3	P_5	P_7	P_9	P_{12}
1	0.471	-1.0	-10.8	-20.5	-25.2	-27.0	-30.0	-30.0
3	0.784	-3.0	-6.9	-16.6	-21.0	-24.0	-27.0	-27.0
6	1.046	-6.0	-5.2	-14.7	-19.2	-22.2	-24.0	-25.2
9	1.208	-9.0	-4.5	-14.1	-18.5	-21.5	-24.0	-25.2
12	1.317	-12.0	-4.2	-13.8	-18.2	-21.0	-23.0	-25.2

Table 2. Sideband relative to total power for sine wave modulation

Carrier suppression, dB	k	P_0	P_1	P_2	P_3	P_4
1	0.667	-1.0	-10.0	-25.2	—	—
3	1.124	-3.0	-6.41	-17.0	-30.0	—
6	1.518	-6.0	-5.0	-12.5	-24.0	—
9	1.774	-9.0	-4.7	-10.5	-20.5	—
12	1.952	-12.0	-4.7	-9.3	-18.2	-30.0

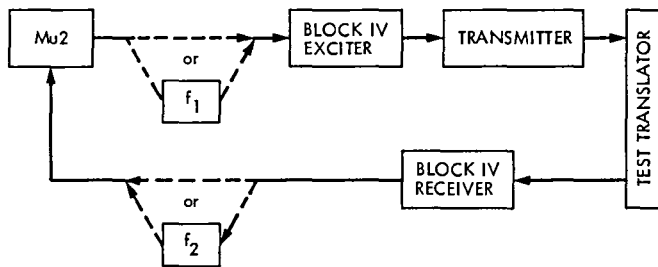


Fig 1. Test configuration for measuring range error

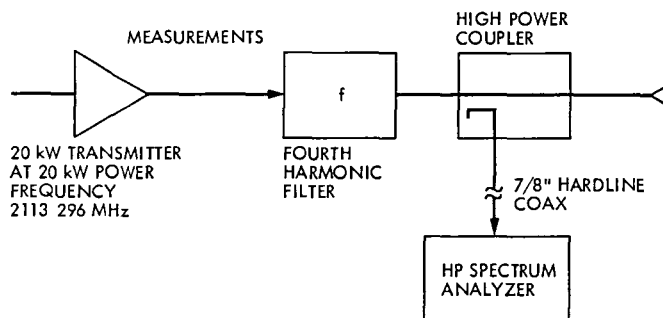


Fig 2. Test configuration for spectral measurements

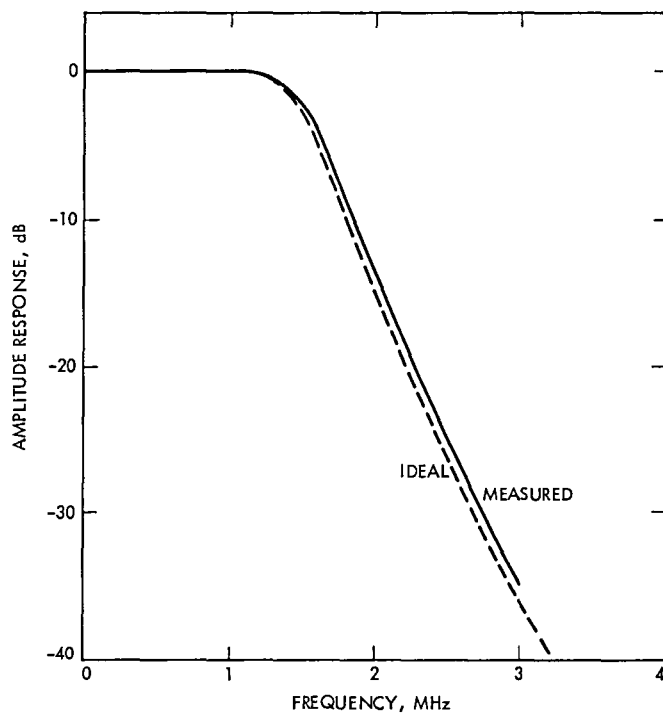
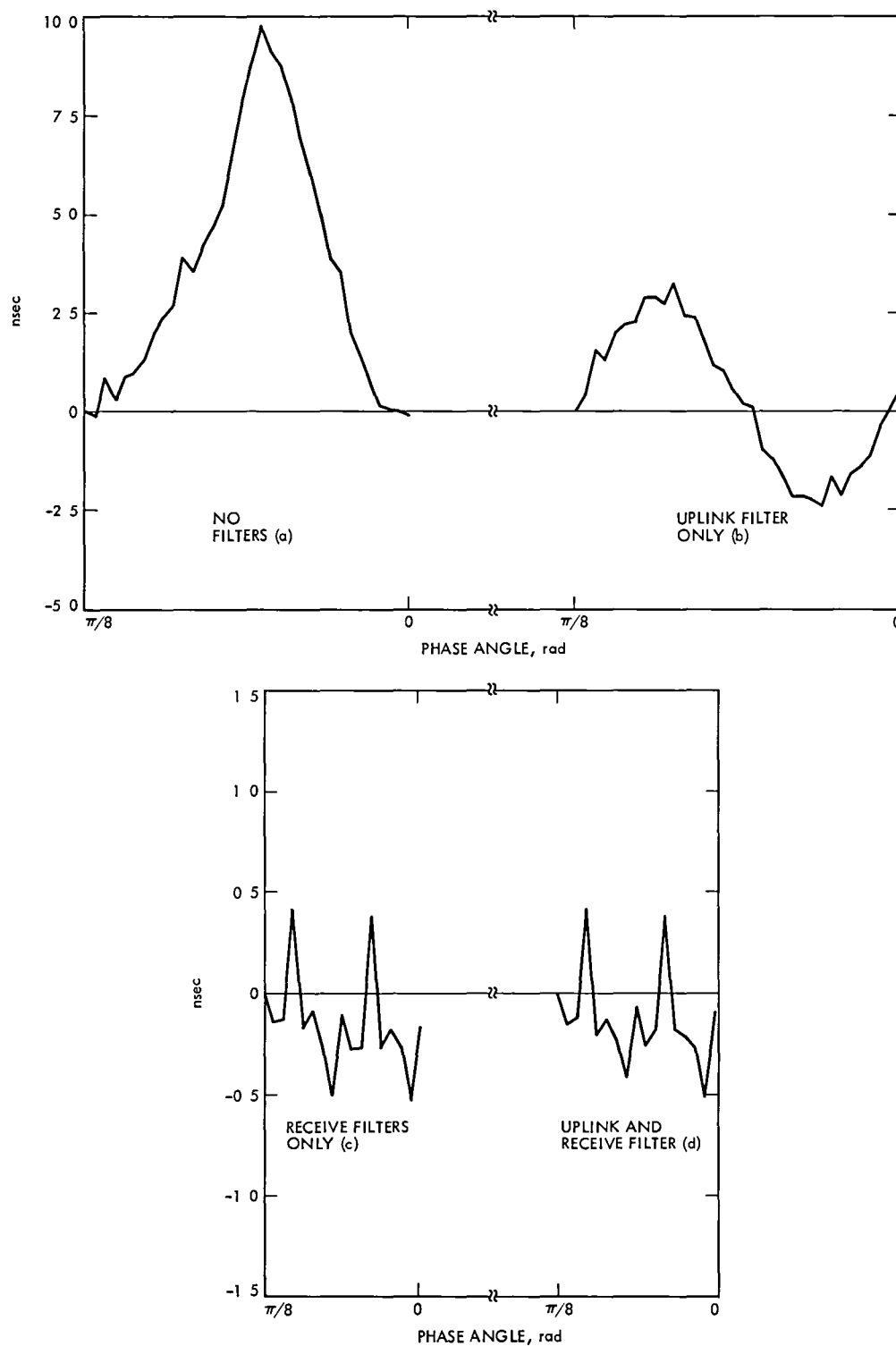
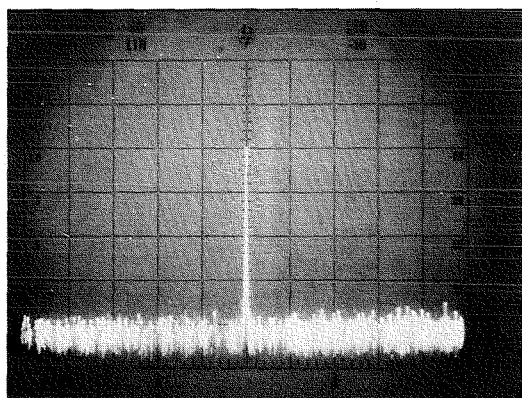


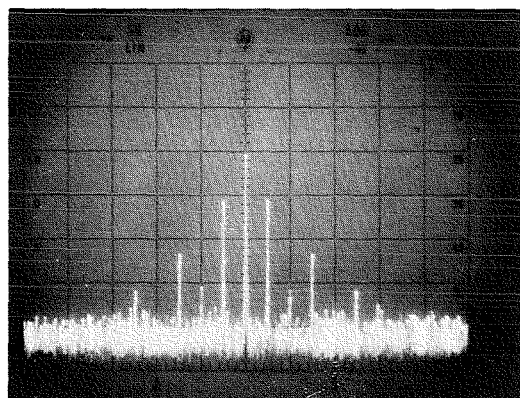
Fig. 3. Response curve for the uplink filter (the filter is a six section Butterworth.)



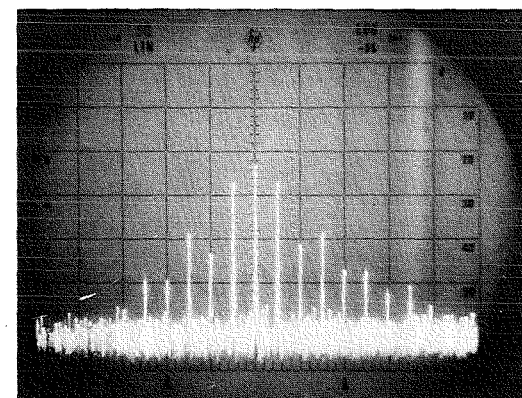
**Fig. 4. Range error for various phase relationships between received and local model range code
(Note that (a) and (b) have a different vertical scale from (c) and (d))**



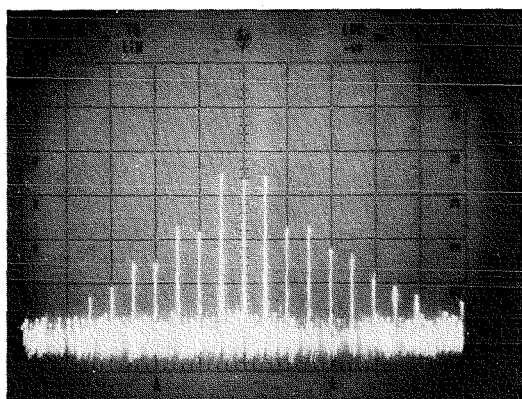
0 dB



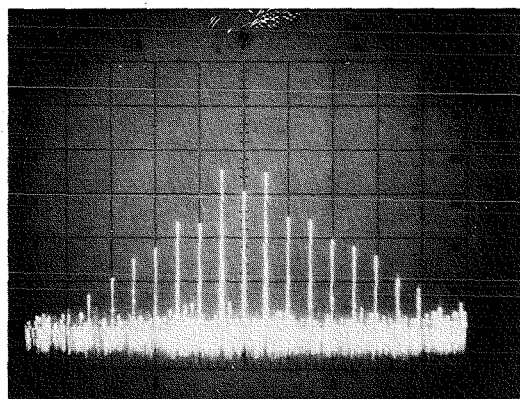
1 dB



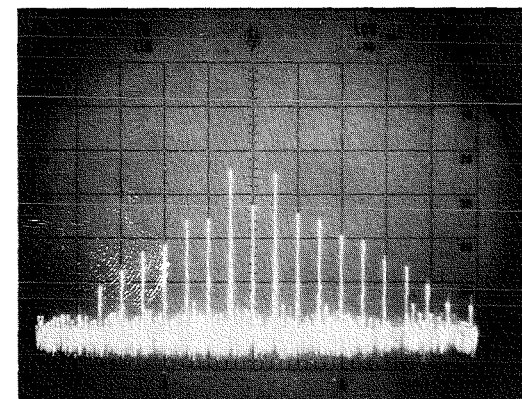
3 dB



6 dB

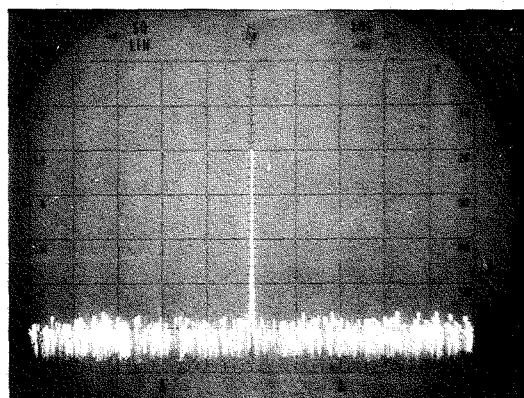


9 dB

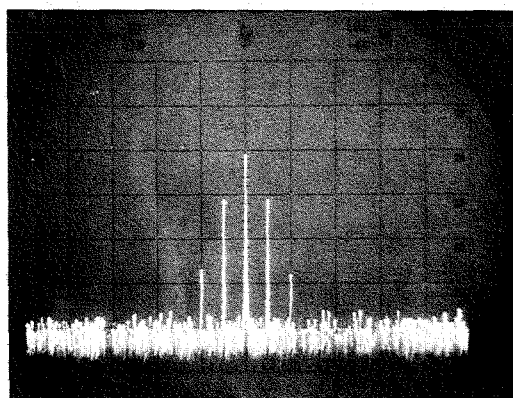


12 dB

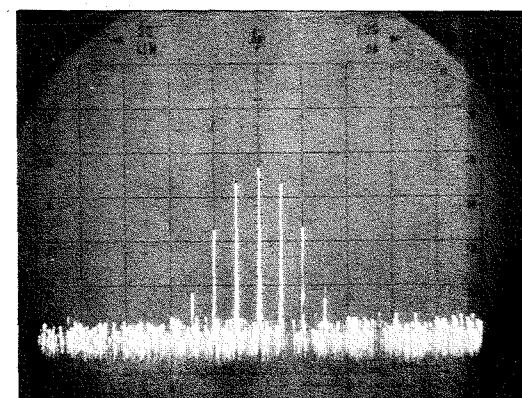
Fig. 5. Spectrum of transmitted signal with no filter; square wave modulation. Numbers indicate carrier suppression



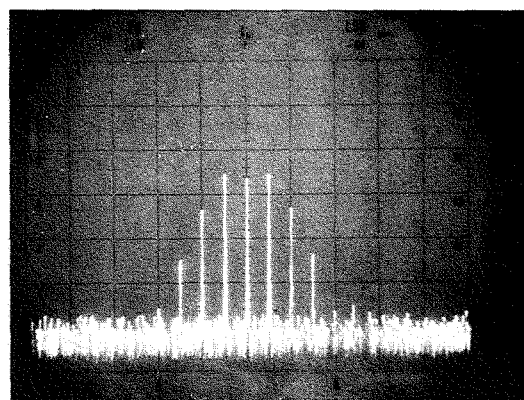
0 dB



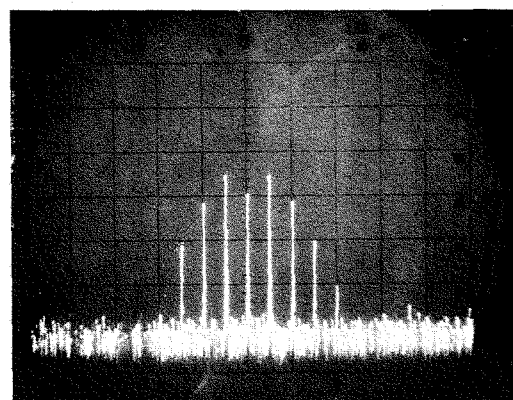
1 dB



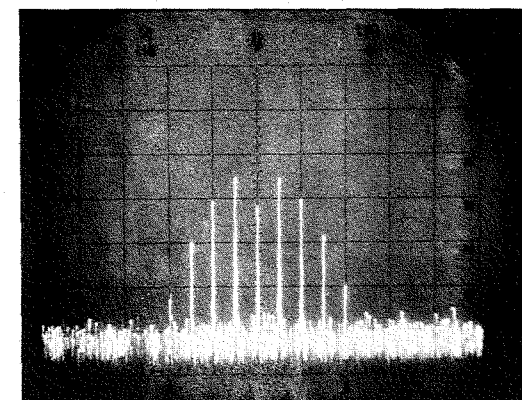
3 dB



6 dB



9 dB



12 dB

Fig. 6. Spectrum of transmitted signal with filter; sine-wave modulation. Numbers indicate carrier suppression.

Radio Frequency Interference by Earth Orbiting Satellites: Deep Space Interference Prediction Program

D L Cain

Tracking Systems and Applications Section

P E Beyer

Control Center Operations

The sensitivity requirements of the DSN Flight Projects make them quite susceptible to radio frequency interference. Efforts to predict this interference from earth orbiting satellites have been ongoing for approximately five years. Software development in the areas of trajectory computations and telecommunications analysis are discussed in this article.

I. Background

Development of the Deep Space Interference Program (DSIP) software began in 1976. Simplified engineering models and trajectory formulae were used. As a result, operational predictions did not meet the accuracy requirements of the DSN. Previous articles (Refs 1 and 2) have defined the response of a ground receiver that is tracking a spacecraft telemetry signal to interference by a close (in frequency) sine wave, or spectral spike signal. A software model was developed to incorporate the receiver response, interface between the standard JPL navigation trajectory program, DPTRAJ, was incorporated in the second level, DSIP2. Hence, DSIP2 has been constructed to compute the times and amount of such interference, given the necessary descriptions of the spacecraft and satellite signals and ephemerides.

II. Problem Statement

Given the definition of what constitutes a state of interference (roughly speaking, this consists of a satellite output spike

being near enough to a received spike — either the carrier signal or a telemetry subcarrier or harmonic), and also given the spacecraft and satellite orbital information, compute common view periods, compute signal strengths and doppler shifts and then finally construct a table of predicted interference indicating spacecraft, satellite, station, bit rate mode and type of interference.

III. Input Required

The interfering signal must be defined. This consists of a list of discrete (spike) frequencies and corresponding output powers. Currently CW is the only type of interfering signal we have a model for (a model for spread spectrum or pulse data is expected to be available in the future). In addition, to compute doppler shifts, orbital data for each satellite must be provided.

The above data is derived from orbital elements or initial conditions typically furnished by outside agencies. Further-

more, the orbital elements sometimes need conversion (for example, mean elements often provided must be converted to osculating elements, i.e., initial conditions required for subsequent integration, this is done by a computational algorithm supplied by the external agency) The resulting orbital elements or cartesian position and velocity are propagated forward in time in the form of a stored-file ephemeris In the past, this ephemeris was provided by DPTRAJ. As part of ongoing program development, an alternate program, SATRAP, was built to reduce costs (described below in Section IV)

Next the spacecraft signal is characterized The transmitter power, modulation index, and subcarrier frequencies are specified, and corresponding bit rates and coding schemes defined These are treated also as sets of spikes of average power (When these spikes are approached by the satellite spikes, degradation may occur) Since the spacecraft-received carrier frequency depends upon the transmitted uplink frequency in the two-way mode, provision is made via a best-lock algorithm to predict what the uplink frequency will be

For the trajectory data needed for the computations of view periods and doppler shifts, a standard DPTRAJ "P" file is provided by the respective project

IV. Program Logic

Read in the names of the Deep Space Stations (DSSs) of the run

Do for each satellite of the run

 Read in the satellite data

 Compute view periods

Enddo

Do for each spacecraft of the run

 Read in the input spacecraft data

 Compute view periods for the spacecraft

 Construct a spacecraft pass transmitter frequency table

 Do for each DSS of the run

 Do for each satellite of the run

 Do for each spacecraft pass

 Compute the spacecraft/satellite view period overlap intervals

 Do for each view period overlap interval

 Look for RFI

 Enddo

 If any RFI occurred during this pass print the spacecraft pass interference tables

 Endif

 Enddo

Enddo

Enddo

Enddo

Stop

V. Program RFI Definitions

Four types of RFI are distinguished by the program

A. Receiver Drop Lock

The receiver is predicted to drop lock (via phase lock jump phenomenon, Ref 2) whenever either (a) the interfering signal is inside the phase lock loop (PLL) band and stronger than the carrier, or (b) the interfering signal is outside the PLL bandwidth and if the ratio of the interfering power to the carrier power is greater than the ratio $\Delta f/B$ where Δf , B are the frequency separation and loop-bandwidth, respectively

A second type of drop-lock occurs when the maser loses enough gain to take the carrier below threshold

B. Receiver Interference

The receiver is adjudged as being interfered with whenever the interference is above threshold (~ 175 dBm) and within one kilohertz of the carrier

C. Telemetry Drop Lock

When conditions similar to Subsection A above occur to a subcarrier, it also suffers drop of lock This of course is a function of the carrier strength, spacecraft output power, mod index, bit rate, and coding scheme

D. Signal-to-Noise Degradation

From theoretical and test data, P Low (Ref 1) developed a model for the decrease in signal-to-noise ratio in a telemetry stream due to an interfering spike near a carrier or subcarrier That model [Eq. 6, p. 206 of Ref. 1] has been incorporated

into DSIP2, except that the $\sin(x)/x$ function has been replaced by a more conservative and easier to compute decreasing-magnitude $(1/n)$ square wave. When by that expression the signal-to-noise ratio drops by more than 0.5 dB, then degradation is predicted to occur.

VI. New Developments

A. DSIP2

The program DSIP2 has been converted to the Univac 1100/81. This involved changing the source code formerly SFTRAN (no longer available) to the newer language SF3. The program was modified to find and interpolate satellite coordinates from a SATRAP file, described next.

B. SATRAP

This special purpose trajectory program was constructed to reduce the high costs of using DPTRAJ, the standard general purpose integrator in use at JPL.

A simple model of spacecraft motion is used (i.e., a non-spherical earth potential, a simplified drag model, radiation pressure, and the sun and moon are the perturbing forces). Only earth satellites can be accommodated. To facilitate the integration, the independent variable is eccentric anomaly rather than time. This change has the characteristic that uniform steps in this variable give smaller steps in time close to earth and longer steps in time farther from earth—desirable in that it is near the earth that the perturbing functions act strongest. Thus, SATRAP uses about twice the step size relative to a time integration (DPTRAJ).

The form of the output file is that of a random access, distributed index file. The concept of volumes, chapters, and verses (a coarse index followed by a finer index, and then one even finer) is used. Thus, with a few file accesses, first inquiring what "volume", then what "chapter," and finally what "verse" is involved (the final one giving the times desired), one

can find records bracketing any given requested time. Another analogy would be a hypothetical congressional record, first bound in volumes by year, then a chapter for each month, and then verses for daily proceedings. Such considerations are important since one satellite can be involved in a hundred orbits in a week.

VII. Future Developments

- (1) There is a need to develop improved interference models for spread-spectrum signals.
- (2) The same is true for pulsed signals.
- (3) Still further development is needed in trajectory computation and/or program organization to reduce costs and increase ease of use.
- (4) The advent of the new Support System computer and its planned teletype interface will result in a more rapid, more automated operation.

VIII. Accuracy of Performance

- (1) Of the nine instances of Voyager I telemetry drop-lock attributable to the Cosmos satellite (from January 1 to October 15, 1981), all were predicted by DSIP2.
- (2) Of the fifteen instances of Pioneer 10 telemetry drop-lock or receiver drop-lock in 1980, eleven were predicted. The other four occurred shortly after launch, hence not enough was known of this particular Cosmos spectrum, since that data must be acquired by observation.

In summary, the whole effort appears successful and with the oncoming increase in the number of S- and X-band satellites and with the high-earth orbiters requiring DSN support (as well as becoming themselves potential interferers) the effort to predict and avoid RFI will continue.

References

1. Low, P. W., "Radio Frequency Interference Effects of Continuous Wave Signals on Telemetry Data. Part II", *The Deep Space Network Progress Report 42-51*, Jet Propulsion Laboratory, Pasadena, California, March and April 1979. See also *ibid*, 42-40, May and June 1977 for first part.
2. Kliger, I. E., and Olenberger, C. F., "Phase-Lock Loop Jump Phenomenon in the Presence of Two Signals", *IEEE Trans on Aerospace and Electronic Systems*, Vol AES-12, No. 1, January 1976, pp. 55-63.

New S-Band Transmitter Automation Software

W Stahnke

Radio Frequency and Microwave Subsystem Section

This article describes the status of the 20-kW S-band transmitter automation project. A new software design using a simplified multitasking approach is described that will improve subsystem performance, maintainability and extensibility.

I. Introduction

Previous efforts toward transmitter automation have culminated in the development of a complete, functionally correct program for operating the 100-kW S-band transmitter. This program was developed over an extended period of time as an SRT effort, and it has provided a great deal of insight into the approaches required to automate a transmitter. Structurally, however, the current program could be strengthened. For implementation into the DSN transmitters, it is proposed that a revised program with multitasking be written that should improve subsystem performance, maintainability and extensibility.

II. Rationale for the Design Approach

From the programmer's point of view, the transmitter controller is a general-purpose microcomputer operating a real-time control system. Like many other such systems, it has the following requirements:

- (1) There must be a means of responding to events occurring asynchronously in the physical environment
- (2) There must be a means of controlling several devices concurrently

What sets the transmitter controller apart from some other systems, however, is that it requires very little computational

work to be done. Almost all computations fall into the category of integer arithmetic.

In addition, the transmitter control function is simplified by two special features:

- (1) Hardware interlocks provide high-speed response to potentially catastrophic events
- (2) Most of the controlled devices are low-speed electro-mechanical components

As a result, only moderate response time and moderate control speed are required.

To fulfill the requirements outlined above, a multiprocessing or multitasking approach is preferred. In this case, a single CPU provides ample processing power, so there is little incentive to introduce multiple processors. Because of the abundance of processing power, it is not necessary to make optimal use of the processor itself, so a conventional preemptive, priority-based real-time multitasking executive is not required. Instead, the objectives can be achieved with a simplified executive that distributes processor time uniformly among the various tasks. This approach brings the advantages of multitasking to the program without the usual complexities of a full multitasking executive.

This simplified executive function can be depicted by the top-level flow chart shown in Fig. 1. After power-on initialization, the first task is executed for a given period of time, typically one millisecond. After that time has expired, execution of the first task is interrupted by a hardware timer, and the machine state is stored in a corner of memory reserved for that purpose. The second task is then executed for an identical period, until it is interrupted. After the last task is interrupted, the machine state is restored in preparation for resuming the first task. This cycle is repeated indefinitely.

It is important to note that in the strictest sense Fig. 1 is not a flow chart at all. Figure 1 implies that (for example) task 2 is initiated when task 1 is finished and relinquishes control of the processor. In fact, task 1 is interrupted by an event external to it, and later resumes execution at the same point at which execution was previously interrupted. The code for task 1 (and, in fact, the code for all of the tasks) is written as an endless loop, so that it is never done.

III. Advantages of Simplified Multitasking

The simplified multitasking approach outlined above has several advantages over the polling-loop approach to control systems, including

- (1) Concurrent execution of several tasks
- (2) Faster response to external events
- (3) Programmatically invisible processor allocation

The last point is so important that it deserves a detailed discussion, because it is the principal reason for using multitasking.

In control systems, there are often some functions to be performed that require very little processor time, but require a great deal of elapsed time to execute. One example is filling a tank. The processor must open a valve and wait until the tank is full before closing it. The elapsed time might be several hours.

If one were to write a conventional sequential program to perform this function, the processor would open the valve and then enter a short testing loop to determine if the valve should be closed (see Fig. 2). This simple program might occupy the processor full time for hours, during which time no other functions could be performed (including operator intervention to abort the filling). This is clearly unacceptable.

The only alteration that can provide better performance using a polling-loop approach is to rewrite the program in such a way that it can be called repetitively by the (higher-level)

polling loop. This usually necessitates the use of state flags to allow the processor to determine what it was doing on the most recent pass. For our simple example, the program might be rewritten as shown in Fig. 3. Here only one state flag is required.

The modified routine provides improved performance at the cost of simplicity and clarity. It is now the programmer's responsibility to break the task into subtasks, each of which must execute quickly enough to prevent undue delay to the polling loop. The program and its flowchart have become more difficult to read.

At the most basic level, the difference between the sequential code and modified sequential code is quite simple. In the former, the state of the system is implied by position in the code; in the latter, it is explicitly given by state flags, which must be defined and manipulated by the programmer. The programmer must also be aware of the timing constraints that each routine imposes on the polling loop.

A completely different solution is possible by introducing multitasking. The program can now be written using purely sequential coding techniques, and the multitasking executive allocates processor time to each task automatically. Each task is written as a free-standing sequential program, without regard to polling time. A task is concerned only with itself; the other tasks are invisible to it. This greatly simplifies program design, coding and debugging.

It should be clear at this point that introducing multitasking does not increase the complexity of a real-time program; on the contrary, it enables the programmer for the first time to write simple, purely sequential routines for each application task.

The simplified multitasking approach presented above also has some advantages over conventional preemptive priority-based multitasking. Among these are

- (1) The executive is simplified
- (2) Fewer interrupts are required

IV. Functions of the Executive

The executive performs two separate functions: initialization and multitasking. It also provides a global real-time clock for use by all of the tasks. These functions are described in greater detail in the following paragraphs.

Initialization by the executive encompasses both hardware and software initialization. Hardware initialization is necessary because on power-up the previous state of the hardware is lost,

and the hardware must be set to a known state before multitasking can occur. Software initialization includes setting the initialization semaphore and initializing the storage area for each task.

The initialization semaphore is a single byte whose value is set after power-up to the number of tasks. Each task has associated with it an initialization routine of its own that may require an indefinite amount of time to execute. After finishing its initialization, each task decrements the initialization semaphore and then enters a testing loop that waits until the value of the semaphore is zero. In this way, all of the tasks complete their initialization before any task begins to execute working code. This is shown in the flowchart of Fig. 4 for a single task.

The storage area associated with each task (Fig. 5) must be partially initialized before task rotation can occur. A minimum of two 16-bit register images must be filled: the program counter image and the stack pointer image. Each program counter image must be set to the address of the first instruction of its corresponding task initialization routine. Each stack pointer image must be set to the ceiling value of the stack area reserved for the executive. This is necessary because task rotation requires the existence of a user stack.

The global real-time clock is maintained in memory, and must be reset to zero after power-up. Subsequently, its value is available to all of the tasks for timing physical events. The clock is incremented every time control returns to the executive from a task. The initialization sequence is shown in Fig. 6.

Task rotation is initiated by a hardware timer interrupt. There is a separate task rotating routine for each task. Task rotation is straightforward. The register contents are stored in the register images of the previous task storage area. After updating the interrupt dispatch and global real-time clock, the registers are restored from the next task storage area. Control is then passed to the next task. This sequence is shown in Fig. 7.

V. Present Status

Design and coding of the 20-kW S-band transmitter controller program is proceeding using the multitasking approach outlined above. The executive has been completed, and the hardware drivers are currently being written. Later reports will cover testing of the controller with the simulator for evaluation and support of unattended operation demonstration.

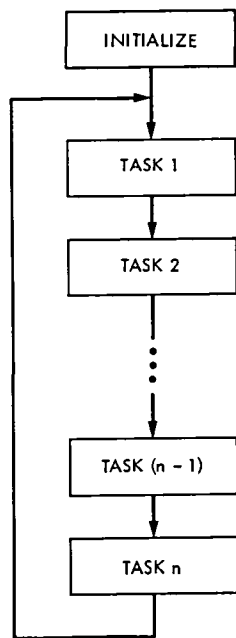


Fig. 1 Executive top-level flowchart

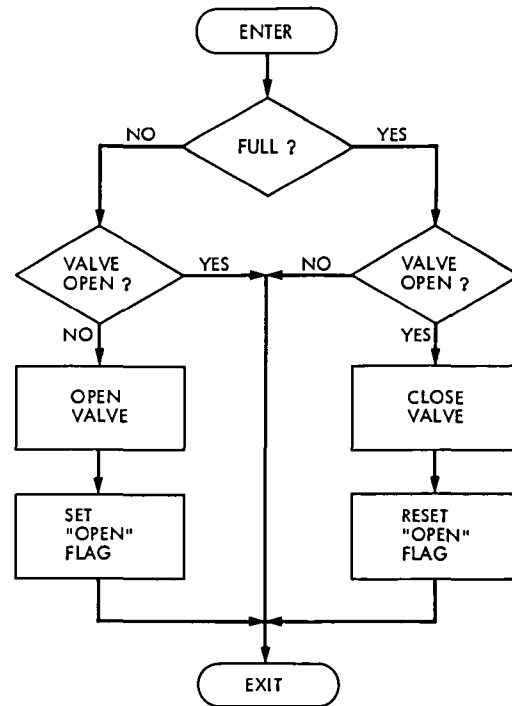


Fig. 3. Valve control flowchart (modified sequential code)

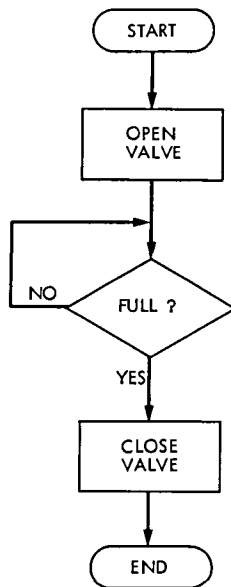


Fig. 2 Valve control flowchart (sequential code)

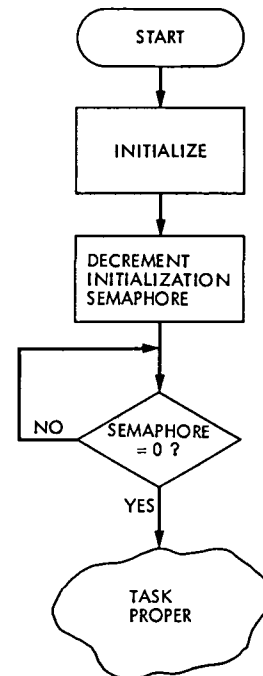


Fig. 4. Task initialization flowchart

H REGISTER	L REGISTER
D REGISTER	E REGISTER
B REGISTER	C REGISTER
ACCUMULATOR	FLAGS
STACK POINTER	
PROGRAM COUNTER	

Fig 5 Register image storage area for a single task

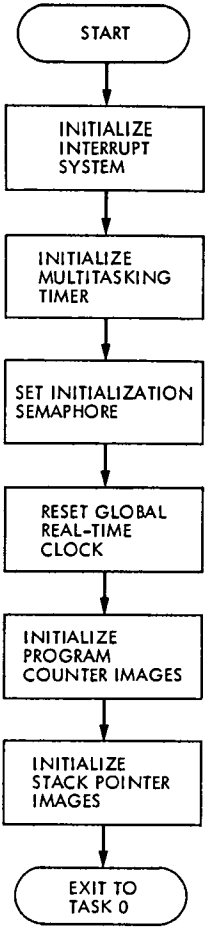


Fig. 6 Executive initialization sequence

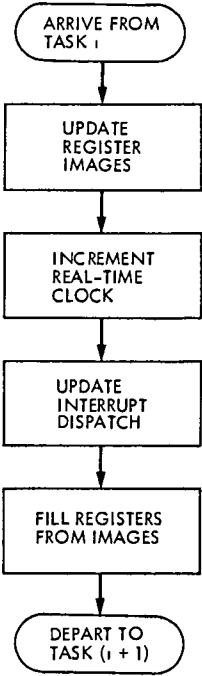


Fig. 7. Task rotation sequence (typical)

The Effect of Partial Coherence in Receiving System Noise Temperature on Array Gain for Telemetry and Radio Frequency Carrier Reception for Similar Receiving Systems

M. H. Brockman

Telecommunications Science and Engineering Division

Signal-to-noise ratio improvement realized by arraying receiving systems or stations for coherent reception is reduced when some portion of predetection noise is coherent in the array. This report examines this effect for arrayed receiving systems with essentially equal apertures including presence of a planet in the array beamwidth.

I. Introduction

In earlier reports "Radio Frequency Carrier Arraying for High-Rate Telemetry" (Ref 1) and "Radio Frequency Carrier Arraying for Near-Maximum Carrier Signal-to-Noise Ratio Improvement" (Ref 2), the various components of operating system noise temperature (which produce the receiver predetection noise) were treated as statistically independent among the receiving systems of an array. This report considers the effect on predetection signal-to-noise ratio improvement for coherent carrier reception and demodulation for the condition when some portion of the predetection noise is not statistically independent but is coherent among the various receiving systems of the array (for some period of time). The following material is directed toward high-rate telemetry reception with residual radio frequency (RF) carrier. Performance shown herein applies to similar receiving systems with equal predetection noise. A later report will consider the case for unequal predetection noise.

II. Receiver Configuration and Predetection Noise

Figure 1 illustrates a method for achieving RF carrier arraying which provides adequate RF carrier signal-to-noise ratio improvement for high-rate telemetry with residual RF carrier. This configuration was presented in Ref 1 with a detailed discussion of the received signal as it passes through the receiving systems. Although only two systems are shown in Fig 1, additional systems can be added to increase the signal-to-noise ratio improvement for high-rate telemetry with residual RF carrier.

Figure 1 can be modified so that much larger antenna separations for the array can be handled conveniently. This modification, described below, represents in general a proposed array configuration for the Mark IV-A DSN system design. In particular, Fig. 1 can be modified so that for receiving system 2, the received signal is down-converted (after the microwave low-noise amplifier) to

an adequately high IF signal frequency that will provide sufficient bandwidth. This down-conversion is accomplished using a stable fixed frequency first local oscillator. The down-converted IF signal is fed over some distance to a mixer located in a central area closely associated with receiving system 1. The local oscillator for this mixer (located in the central area) is effectively the first local oscillator for receiving system 1 down-converted to the appropriate frequency. The resultant output from this mixer (centrally located) is then fed into the IF amplifier (with AGC), shown in Fig 1. Receiving systems 3 through N are also configured in a manner similar to receiving system 2.

As described in Ref 1, the predetection noise in each receiving system is measured relative to reference temperature load(s) connected to the low noise amplifier input (during the measurement) and designated as due to the operating equivalent system noise temperature T_{op1} , T_{op2} , etc. This predetection noise represents a combination of galactic noise, planetary radiation, atmospheric noise, noise in the antenna sidelobes due to the Earth, noise due to losses in microwave reflectors, and noise due to losses in microwave components all lumped with noise due to the input amplifier(s). The following section will develop an expression for predetection signal-to-noise ratio improvement and the resultant phase noise on the first local oscillator for the case where a portion of the receiver noise is coherent among the various receiving systems of the array. It should be noted that for the results presented in this report, the receiving systems are considered to be similar to each other with essentially equal size antenna apertures and/or equal predetection noise (including the portion which is coherent) for the receiving systems of the array.

III. Predetection Signal-to-Noise Ratio and Resultant Carrier Tracking Loop Phase Noise

With the other receiving system(s) (2 through N) switched out of the summing junction (see Fig 1), the predetection carrier signal-to-noise power ratio in receiving system 1 is (from Ref 1)

$$\frac{P_{c1}}{P_{n1}} = \frac{A_1 \cos m_{pd}^2}{NBW_{F_{A1}} N_{o1}} \quad \text{or} \quad \frac{P_{c1}}{NBW_{F_{A1}} N_{o1}} \quad (1)$$

where m_{pd} is the peak phase modulation index and $NBW_{F_{A1}}$ represents the noise bandwidth of the second IF filter F_{A1} . The term N_{o1} is the one-sided noise spectral density for receiver 1 related to T_{op1} as discussed in the preceding section of this report and in Ref 1. The receiving system contains a second-order RF carrier phase tracking loop which utilizes a bandpass limiter and a sinusoidal phase detector. The resultant rms phase noise σ_{ϕ_n} at the output of the RF carrier tracking loop (i.e., on the first local oscillator) due to the predetection carrier signal-to-noise ratio within the closed-loop noise bandwidth of the RF carrier tracking loop is (from Ref 1)

$$\sigma_{\phi_{n1}} = \frac{\frac{N_{o1} \cdot 2B_{L1}}{2}}{P_{c1}} \left[\frac{1 + \frac{P_{c1}}{NBW_{F_{A1}} \cdot N_{o1}}}{0.862 + \frac{P_{c1}}{NBW_{F_{A1}} \cdot N_{o1}}} \cdot \frac{\exp\left(\frac{N_{o1}B_{L1}}{P_{c1}}\right)}{\sinh\left(\frac{N_{o1}B_{L1}}{P_{c1}}\right)} \right]^{1/2} \quad \text{rad, rms} \quad (2)$$

where two-sided closed loop noise bandwidth $2B_{L1}$ can be expressed as

$$2B_{L1} = \frac{2B_{Lo1}}{r_o + 1} \left(1 + r_o \frac{\alpha}{\alpha_{o1}} \right)$$

The term r_o is equal to 2 and α is the bandpass limiter suppression factor (Ref 1)

Next consider receiving system 2 where a portion of its predetection noise is coherent with a similar portion of the predetection noise in receiving system 1 at the output of the summing junction. Designate ϵ as that portion of the predetection

noise that is coherent, then $1 - \epsilon$ is that portion which is statistically independent. With receiving system 2 connected to the summing junction, the summed carrier predetection signal-to-noise ratio at the output of the summing junction becomes

$$\frac{P_{c1\Sigma 1,2}}{P_{n1\Sigma 1,2}} = \frac{(A_1 \cos m_{pd} + \beta_2 A_2 \cos m_{pd})^2}{\left[NBW_{F_{A1}} (1 - \epsilon_1) N_{o1} + \beta_2^2 NBW_{F_{A2}} (1 - \epsilon_2) N_{o2} \right] + \left[\left(NBW_{F_{A1}} \cdot \epsilon_1 N_{o1} \right)^{1/2} + \beta_2 \left(NBW_{F_{A2}} \cdot \epsilon_2 N_{o2} \right)^{1/2} \right]^2} \quad (3)$$

For similar systems with essentially equal diameter antennas and aperture efficiencies, and/or equal predetection noise spectral density (including the portion which is coherent), expression (3) can be written as

$$\frac{P_{c1\Sigma 1,2}}{P_{n1\Sigma 1,2}} = \frac{P_{c1}}{NBW_{F_{A1}} \cdot N_{o1}} \cdot \frac{(1 + \beta_2 \gamma_2)^2}{(1 - \epsilon)(1 + \beta_2^2) + \epsilon(1 + \beta_2)^2} \quad (4)$$

where β_2 is the voltage coupling of receiving system 2 relative to receiving system 1 at the output of the summing junction, γ_2^2 is the carrier power-to-noise spectral density ratio of receiving system 2 relative to receiving system 1 and $\epsilon = \epsilon_1 = \epsilon_2$. For N receiving systems arrayed, the predetection carrier signal-to-noise ratio in receiving system 1 is

$$\frac{P_{c1\Sigma 1, \dots, N}}{P_{n1\Sigma 1, \dots, N}} = \frac{(A_1 \cos m_{pd} + \beta_2 A_2 \cos m_{pd} + \dots + \beta_N A_N \cos m_{pd})^2}{\left[NBW_{F_{A1}} (1 - \epsilon_1) N_{o1} + \beta_2^2 NBW_{F_{A2}} (1 - \epsilon_2) N_{o2} + \dots + \beta_N^2 NBW_{F_{AN}} (1 - \epsilon_N) N_{oN} \right] + \left[\left(NBW_{F_{A1}} \epsilon_1 N_{o1} \right)^{1/2} + \beta_2 \left(NBW_{F_{A2}} \epsilon_2 N_{o2} \right)^{1/2} + \dots + \beta_N \left(NBW_{F_{AN}} \epsilon_N N_{oN} \right)^{1/2} \right]^2} \quad (5)$$

For similar systems (as described above for expression 3), expression (5) becomes

$$\frac{P_{c1\Sigma 1, \dots, N}}{P_{n1\Sigma 1, \dots, N}} = \frac{P_{c1}}{NBW_{F_{A1}} \cdot N_{o1}} \cdot \frac{(1 + \beta_2 \gamma_2 + \dots + \beta_N \gamma_N)^2}{(1 - \epsilon)(1 + \beta_2^2 + \dots + \beta_N^2) + \epsilon(1 + \beta_2 + \dots + \beta_N)^2} \quad (6)$$

The resultant rms phase noise ($\sigma_{\phi_{n1\Sigma 1,2}}$) at the output of the RF carrier tracking loop (i.e., on the first local oscillator signal) in receiving system 1 due to the predetection signal-to-noise ratio within the closed-loop noise bandwidth of the RF carrier tracking loop becomes

$$\sigma_{\phi_{n1\Sigma 1,2}} = \frac{\frac{N_{o1}}{2} \cdot 2B_{L1}}{P_{c1}} \cdot \frac{1}{n_2'} \left[\frac{1 + \frac{P_{c1} \eta_2'}{NBW_{F_{A1}} \cdot N_{o1}}}{0.862 + \frac{P_{c1} \cdot \eta_2'}{NBW_{F_{A1}} \cdot N_{o1}}} \cdot \frac{\exp\left(\frac{N_{o1} \cdot B_{L1}}{P_{c1} \cdot \eta_2'}\right)}{\sinh\left(\frac{N_{o1} \cdot B_{L1}}{P_{c1} \cdot \eta_2'}\right)} \right]^{1/2} \text{ rad, rms} \quad (7)$$

where

$$\eta_2' = \frac{(1 + \beta_2 \gamma_2)^2}{(1 - \epsilon)(1 + \beta_2^2) + \epsilon(1 + \beta_2)^2} \quad (7a)$$

for two similar receiving systems arrayed. For N similar receiving systems arrayed, the resultant rms phase noise ($\sigma_{\phi_{n1\Sigma 1, \dots, N}}$) becomes

$$\sigma_{\phi_{n1\Sigma 1, \dots, N}} = \frac{\frac{N_{oN} \cdot 2B_{L1}}{2}}{P_{c1}} \cdot \frac{1}{\eta'_N} \left[\frac{1 + \frac{P_{c1} \cdot \eta'_N}{NBW_{FA1} \cdot N_{o1}}}{0.862 + \frac{P_{c1} \cdot \eta'_N}{NBW_{FA1} \cdot N_{o1}}} \cdot \frac{\exp\left(\frac{N_{o1} \cdot B_{L1}}{P_{c1} \cdot \eta'_N}\right)}{\sinh\left(\frac{N_{o1} \cdot B_{L1}}{P_{c1} \cdot \eta'_N}\right)} \right]^{1/2} \text{ rad, rms} \quad (8)$$

where

$$\eta'_N = \frac{(1 + \beta_2 \gamma_2 + \dots + \beta_N \gamma_N)^2}{(1 - \epsilon)(1 + \beta_2^2 + \dots + \beta_N^2) + \epsilon(1 + \beta_2 + \dots + \beta_N)^2} \quad (8a)$$

The discussion above has addressed predetection carrier signal-to-noise ratio improvement and resultant phase noise. It should be noted that for the condition where the varying group delay on the telemetry sidebands is tracked out among the various receiving systems of the array (which is accomplished at baseband for the configuration shown in Fig. 1), the improvement in telemetry signal-to-noise ratio is also represented by expressions (7a) and (8a) for the condition where the telemetry signal combiner has an assumed zero loss and the loss due to carrier tracking loop phase noise (radio loss) is negligible.

Up to this point, phase noise on the first local oscillator due to predetection carrier signal-to-noise power ratio has been examined. Total phase noise on the first local oscillator also includes phase noise from the local oscillator (VCO \times Q) of receiving system 2 through N , which is coupled through the summing junction (as developed in Ref. 1). Consequently, the total phase noise at the output of the principal carrier tracking loop (i.e., on the first local oscillator) becomes

$$\left[\sigma_{\phi_{n1\Sigma 1,2}}^2 + \left(\frac{\beta_2 \sigma_{\phi_{n2}}}{1 + \beta_2} \right)^2 \right]^{1/2} \quad (9)$$

for two receiving systems arrayed. For N receiving systems arrayed, the total rms phase noise on the first local oscillator is

$$\left[\sigma_{\phi_{n1\Sigma, \dots, N}}^2 + \left(\frac{\beta_2 \sigma_{\phi_{n2}}}{1 + \beta_2} \right)^2 + \dots + \left(\frac{\beta_N \sigma_{\phi_{nN}}}{1 + \beta_N} \right)^2 \right]^{1/2} \quad (10)$$

As developed in Ref. 1, the total rms phase noise in expressions (9) and (10) can be considered as due to an equivalent predetection carrier signal-to-noise ratio within the closed-loop noise bandwidth of the RF carrier phase tracking loop. Comparison of this equivalent carrier signal-to-noise ratio in a single receiving system (i.e., system 1) provides the improvement due to radio frequency carrier arraying for the high-rate telemetry configuration where some portion of the predetection noise is not statistically independent among the various receiving systems of the array.

As described in Ref. 1, the RF carrier phase tracking loops in receiving systems 2 through N are also second-order loops which utilize a bandpass limiter and sinusoidal phase detector. Since the closed-loop noise bandwidth of the carrier phase tracking loop for receiving systems 2 through N is much narrower (by design) than that in receiving system 1, phase noise in receiving system 1 carrier tracking loop produces a reduction in predetection signal-to-noise ratio for receiving systems 2 through N . The resultant predetection carrier signal-to-noise ratio in receiving system 2 for two systems arrayed is then

$$\frac{P_{c2\Sigma 1,2}}{P_{n2}} = \frac{P_{c2}}{NBW_{FA2} \cdot N_{o2}} \left(1 - \frac{\sigma_{\phi_{n1\Sigma 1,2}}^2}{2} \right)^2 \quad (11)$$

which produces an rms phase noise

$$\sigma_{\phi_{n2\Sigma 1,2}} = \frac{\frac{N_{o2}}{2} \cdot 2B_{L2}}{P_{c2\Sigma 1,2}} \cdot \left[\frac{1 + \frac{P_{c2\Sigma 1,2}}{NBW_{FA2} \cdot N_{o2}}}{0.862 + \frac{P_{c2\Sigma 1,2}}{NBW_{FA2} \cdot N_{o2}}} \cdot \frac{\exp\left(\frac{N_{o2} \cdot B_{L2}}{P_{c2\Sigma 1,2}}\right)}{\sinh\left(\frac{N_{o2} \cdot B_{L2}}{P_{c2\Sigma 1,2}}\right)} \right]^{1/4} \text{ rad, rms} \quad (12)$$

For N systems arrayed, the predetection carrier signal-to-noise ratio in receiving system 2 is

$$\frac{P_{c2\Sigma 1, \dots, N}}{P_{n2}} = \frac{P_{c2} \left(1 - \frac{\sigma_{\phi_{n1\Sigma 1, \dots, N}}^2}{2}\right)^2}{NBW_{FA2} \cdot N_{o2}} \quad (13)$$

and the predetection carrier signal-to-noise ratio in receiving system N is

$$\frac{P_{cN\Sigma 1, \dots, N}}{P_{nN}} = \frac{P_{cN} \left(1 - \frac{\sigma_{\phi_{n1\Sigma, \dots, N}}^2}{2}\right)^2}{NBW_{FAN} \cdot N_{oN}} \quad (14)$$

Substitution of corresponding terms from (13) or (14) into expression (12) provides $\sigma_{\phi_{n2\Sigma 1, \dots, N}}$ or $\sigma_{\phi_{N\Sigma 1, \dots, N}}$ for receiving system 2 or N respectively (Ref 1). As noted previously (Ref 1), the term $\sigma_{\phi_{n1\Sigma 1, \dots, N}}$ in expression (13) should be replaced by an rms value similar to that shown in expression (10) with the term $\beta_2 \sigma_{\phi_{n2}} / (1 + \beta_2)$ deleted. This iteration results in a very small change in carrier signal-to-noise ratio improvement for the parameters considered here. A similar iteration applies to expression (14).

IV. Performance

Improvement in predetection signal-to-noise ratio can be determined from expressions (7a) and (8a) developed in Section III of this article. These expressions apply under the condition that in a predetection bandwidth which encompasses the signal of interest, the phase shift and group delay in the various receiving systems of the array are essentially matched. In addition, as outlined in Section III, the varying group delay on the telemetry sidebands is tracked out among the various receiving systems of the array by the telemetry signal combiner during a station pass. Also, for telemetry, the effective signal-to-noise ratio improvement obtained does not include the telemetry signal combiner loss and the loss (radio loss) due to carrier tracking loop phase noise (σ_{ϕ_n}).

Figure 2 shows the improvement in predetection signal-to-noise ratio as a function of the statistically independent portion of the predetection noise ($1 - \epsilon$) with two receiving systems arrayed for γ_2 values of 1.0, 0.95, 0.90, and 0.84 for the case where $\beta_2 = \gamma_2$. As defined in Section III, γ_2^2 is the carrier power-to-noise spectral density ratio of receiving system 2 relative to receiving system 1, and β_2 is the voltage coupling of receiving system 2 relative to receiving system 1 at the summing junction. Improvement in RF carrier predetection signal-to-noise ratio is maximum for $\beta = \gamma$ when the predetection noise is all statistically independent (noncoherent) or $1 - \epsilon = 1$, which is generally representative for the DSN. This maximum has a broad peak with variation of β about γ . Figure 2 also represents the effective telemetry signal-to-noise ratio improvement for the conditions described in the preceding paragraph for the case where the voltage coupling (β) = γ in the telemetry signal combining process. The resultant decrease in signal-to-noise ratio improvement is shown in Fig 2 as an increasing portion of predetection noise (ϵ) becomes coherent or as the statistically independent portion of predetection noise ($1 - \epsilon$) decreases. Figures 3, 4, and 5 show similar characteristics for three, four, and six similar receiving systems arrayed. Figures 6, 7 and 8 show the effect on predetection signal-to-noise ratio improvement with γ values of 1.0, 0.95, 0.90 and 0.84 for two, three and four receiving systems arrayed,

respectively, when the summing junction voltage coupling β is varied. Also shown in Figs 6, 7, and 8 is the effect on predetection signal-to-noise ratio improvement when the portion of predetection noise that is statistically independent is 1.0, 0.8 or 0.5

Consider next the equivalent RF carrier signal-to-noise ratio improvement as determined from the total rms phase noise on the first local oscillator (see expressions 9 and 10 and associated discussion in Section III). The following sets of design parameters apply for the performance presented in this report. The sets of parameters for receiving system 1 are

Threshold two-sided noise bandwidth			
$2B_{L_{01}}$	12	152	30 Hz
Predetection IF filter noise bandwidth			
$NBW_{F_{A1}}$	2200	2200	2000 Hz
while the corresponding sets of parameters for receiving system 2 through N are			
Threshold two-sided noise bandwidth			
$2B_{L_{02, \dots, N}}$	0.1	1.0	0.3 Hz
Predetection IF filter noise bandwidth			
$NBW_{F_{A2, \dots, N}}$	2200	2200	2000 Hz

Using the parameters above, Fig 9 shows the equivalent RF carrier signal-to-noise ratio improvement for two similar receiving systems arrayed (high-rate telemetry configuration) as a function of summing junction voltage coupling β_2 . Performance is shown for γ_2 values of 1.0, 0.95, 0.90, and 0.84 with $1 - \epsilon$ values of 1.0, 0.8, and 0.5. Performance for $1 - \epsilon = 1$ represents the same as that shown in Fig 3 (Ref 1) for two systems. Note that, in general, performance variation with voltage coupling β_2 is quite gradual. The information shown in Fig 9 can be rearranged to show RF carrier signal-to-noise ratio improvement as a function of $1 - \epsilon$. Figures 10a and 10b show this characteristic for two similar systems arrayed for the same values of γ_2 as in Fig 9, with voltage coupling values of 1.0 and 0.5, respectively.

Equivalent RF carrier signal-to-noise ratio improvement for three similar receiving systems arrayed (as determined from total rms phase noise on the first local oscillator) is shown in Figs 11, 12a, and 12b as a function of voltage coupling and $1 - \epsilon$, respectively. Similar characteristics are shown in Figs 13, 14a and 14b for four similar receiving systems arrayed while Figs 15, 16a, and 16b apply to six similar receiving systems arrayed.

Some initial measurements of radio-frequency carrier signal-to-noise ratio improvement have been made in the laboratory for two and three receiving systems arrayed. These initial measurements were made with one of the sets of parameters described above. In particular, the predetection IF filter noise bandwidth was 2200 Hz with a $2B_{L_0}$ of 152 Hz for receiving system 1 and a $2B_{L_0}$ of 1 Hz for receiving systems 2 and 3. Measurement with two receiving systems arrayed with $1 - \epsilon_2$ and γ_2 equal to 1 provided an RF carrier signal-to-noise ratio improvement of 2.8 dB for $\beta_2 = 1$, an improvement of 2.4 dB for $\beta_2 = 0.5$, and an improvement of 2.5 dB for $\beta_2 = 1.0$ and $\gamma_2 = 0.94$. The corresponding calculated improvement is 2.9, 2.5 and 2.6 dB respectively. Measurement with $1 - \epsilon_2 = 0.75$ and $\gamma_2 = \beta_2 = 1$ provided a 2.0-dB RF carrier signal-to-noise ratio improvement. An improvement of 1.8 dB was measured for $1 - \epsilon = 0.77$ with $\gamma_2 = 1$ and $\beta_2 = 0.5$. These values agree with predicted performance (see Figs 9, 10a, and 10b).

Measurement with three receiving systems arrayed with $(1 - \epsilon_2) = (1 - \epsilon_3) = 1$ and $\beta_2 = \beta_3 = 1$ provided an RF carrier signal-to-noise ratio improvement of 4.3 dB for $\gamma_2 = \gamma_3 = 1$ and an improvement of 4 dB for $\gamma_2 = 0.94$ and $\gamma_3 = 0.93$, which agree with predicted performance (see Fig 11). Measurement with $(1 - \epsilon_2) = (1 - \epsilon_3) = 0.77$ and $\gamma_2 = \gamma_3 = 1$ provided an RF carrier signal-to-noise ratio improvement of 2.7 dB with $\beta_2 = \beta_3 = 1$ and also with $\beta_2 = \beta_3 = 0.5$. These values agree with predicted performance (see Figs 12a and 12b).

V. Discussion

The following material examines the situation where a planet in the solar system is within the beamwidth of the array considered in this report

Consider the case of two receiving systems (1 and 2) arrayed for coherent reception of a signal from a spacecraft at planetary distance. A planet in the solar system is within the beamwidth of the two antennas of the array. This two-aperture (antenna) array will see the hot body (noise) radiation from the planet as an interferometer operating at a reception frequency wavelength λ . Designate the effective distance between the antennas of two stations as $B_{1,2}$ with $B_{r1,2}$ as the component that is perpendicular to the line of sight to the planet. The fringe spacing of the interferometer is then $\lambda/B_{r1,2}$ radians.

Consider for the moment, a situation where the planet appears as a point noise source. That is, the angular size of the planet as seen from Earth is very small compared to the fringe spacing ($\lambda/B_{r1,2}$). Within the fringe spacing (as the Earth rotates), the noise radiation from the planet as seen in the predetection bandwidth of the array is in phase for the two receivers at the center of an angular segment $\lambda/2B_{r1,2}$, orthogonal at each edge of the angular segment, and 180° out of phase at the center of the adjacent angular segments $\lambda/2B_{r1,2}$ due to continuing change in path length to the two antennas. This in-phase/out-of-phase situation continues alternately during a station pass as the array tracks the spacecraft and it is superimposed on the statistically independent portion of predetection noise. For this situation of a planet appearing as a point noise source and representing a given fractional part (ϵ) of operating equivalent system noise temperature (T_{op}), the decrease in predetection signal-to-noise ratio improvement shown in Fig 2 (for the corresponding $(1 - \epsilon)$) corresponds to the in-phase predetection noise condition at the center of an angular segment $\lambda/2B_{r1,2}$. The orthogonal predetection point noise source condition corresponds to a signal-to-noise ratio improvement for $(1 - \epsilon) = 1$, while the 180° out-of-phase point noise source condition essentially provides signal-to-noise ratio improvement for $(1 - \epsilon) = 1$, but at a T_{op} without the point noise source contribution. The discussion above also applies to RF carrier signal-to-noise ratio improvement as shown in Figs 9, 10a, and 10b for two similar receiving systems arrayed.

A planet of the solar system may not appear as a point noise source for the array considered here. In fact, it can have an angular size that is larger than the fringe spacing $\lambda/B_{r1,2}$. The following development addresses this consideration. In general for the application presented in this report, the planets (as viewed from Earth) can be characterized as disks with a uniform brightness distribution D at the reception frequency. The brightness distribution will be a function of the reception frequency. Utilizing information in Ref. 3 and writing the brightness transform as a function of the angular size (ν) of the planet and the inverse of fringe spacing, the brightness transform R becomes

$$R_{1,2} = \frac{1}{\nu} \int_{-\nu/2}^{+\nu/2} D(\omega) e^{-2\pi i \left(\frac{B_{r1,2}}{\lambda} \right) x} d\chi \quad (15)$$

with integration from the centerpoint of the planet disk out to the edges. Integration of expression (15) provides

$$R_{1,2} = D(\omega) \cdot \frac{\sin \left(\pi \frac{B_{r1,2}}{\lambda} \nu \right)}{\pi \frac{B_{r1,2}}{\lambda} \nu} \quad (16)$$

For the situation where the planet appears as a point noise source, the $\sin x/x$ type of expression above approaches one. The brightness transform becomes the total power of the noise source (planet) as seen in the predetection bandwidth and the preceding discussion relating to a point noise source applies. Figure 17 shows expression (16) plotted as a function of the ratio of planet angular size to fringe spacing $[\nu/(\lambda/B_{r1,2})]$. Note that the above integral is centered ($\nu = 0$) on the in-phase predetection noise situation at the center of an angular segment $\lambda/2B_{r1,2}$ that was discussed earlier in this section of the report. Figure 17 is the resultant amplitude of the brightness transform for this particular path length situation to the two antennas for various ratios of planet angular size to fringe spacing. This figure shows the magnitude of the correlated portion of noise power relative to total

noise power from the planet which is defined as fringe visibility for radio interferometry measurements (Ref 3). Consequently Fig 17 provides the information necessary to scale the coherent portion of predetection noise relative to a point noise source for any ratio of planet angular size to fringe spacing $[\nu/(\lambda/B_{r1,2})]$. Any point on Fig 17 can be treated as an equivalent reduced-in-magnitude point noise source representing a reduced ϵ , and the resulting reduction in predetection and RF carrier signal-to-noise ratio improvement can be determined from Figs 2, 9, 10a and 10b

Consider next three receiving systems (1, 2 and 3) arrayed for coherent reception of a spacecraft signal with a planet within the beamwidth of the three antennas. For the moment, consider a situation where the planet appears as a point noise source which represents a given fractional part (ϵ) of the operating equivalent system noise temperature (T_{op}). At those times when the noise radiation from the planet as seen in the predetection bandwidth of the array is in-phase from receivers 1 and 3 in an angular segment $\lambda/2B_{r1,3}$ at essentially the same time as it is from receivers 1 and 2 in an angular segment $\lambda/2B_{r1,2}$, the decrease in predetection signal-to-noise ratio improvement shown in Fig 3 applies for the corresponding $(1 - \epsilon)$ value. At all other times, the decrease in predetection signal-to-noise ratio improvement will be less. This discussion also applies to RF carrier signal-to-noise ratio improvement shown in Figs. 11, 12a, and 12b for three similar systems arrayed

For the situation where the planet is not a point source, the normalized brightness transforms for three receiving systems become

$$R_{1,2,3} = \frac{D(\omega)}{2} \left(\frac{\sin \left(\pi \frac{B_{r1,2}}{\lambda} \nu \right)}{\pi \frac{B_{r1,2}}{\lambda} \nu} + \frac{\sin \left(\pi \frac{B_{r1,3}}{\lambda} \nu \right)}{\pi \frac{B_{r1,3}}{\lambda} \nu} \right) \quad (17)$$

Note again in this case, the above expression (17) represents a result centered ($\nu = 0$) on the in-phase predetection noise situation at the center of angular segment $\lambda/2B_{r1,2}$ and simultaneously at the center of angular segment $\lambda/2B_{r1,3}$. For this particular path length situation and given fringe spacings $\lambda/B_{r1,2}$ and $\lambda/B_{r1,3}$, expression (17) can be evaluated and the effect of the planet on decrease in predetection and RF signal-to-noise ratio improvement can be determined from Figs 3, 11, 12a, and 12b. Extension of the above discussion can be applied to four or more similar receiving systems arrayed

To illustrate the above, consider a 300-meter $B_{r1,2}$ between two stations arrayed (34-meter-diameter antennas) and operating at a reception frequency of 8420 MHz ($\lambda = 3.56$ cm). The resultant fringe spacing is 24.5 arcsec. The planet Venus is in the beamwidth of the antennas, its contribution to system noise temperature is 10 kelvins and it subtends an angle of 42 arcsec. The ratio 42/24.5 equals 1.7 which, from Fig 17, provides a reduction in relative magnitude of correlated noise power to 0.125 or 1.25 kelvins. The operating system temperature is $20 + 10 = 30$ kelvins. The statistically independent portion $(1 - \epsilon)$ of predetection noise power is $(30 - 1.25)/30$ or 0.958. Reduction in predetection and RF carrier signal-to-noise ratio improvement is 0.16 dB (Figs 2 and 10a) at those times during a station pass when the fractional noise (0.125) from Venus is in-phase in the predetection bandwidth of the two arrayed receiving systems. Note that if the 10 kelvins represented a point noise source, reduction in signal-to-noise ratio improvement would be 1.2 dB instead of 0.16 dB since $1 - \epsilon$ becomes $(30 - 10)/30$ or 0.667.

References

- 1 Brockman, M. H , "Radio Frequency Carrier Arraying for High Rate Telemetry," in *The Deep Space Network Progress Report 42-45*, pp. 209-223, Jet Propulsion Laboratory, Pasadena, Calif , June 15, 1978
- 2 Brockman, M H , "Radio Frequency Carrier Arraying for Near Maximum Carrier Signal-to-Noise Ratio Improvement," in *The Deep Space Network Progress Report 42-49*, pp 99-106, Jet Propulsion Laboratory, Pasadena, Calif , Feb 15, 1979
3. Thomas, J B , An Analysis of Source Structure Effects in Radio Interferometry Measurements, Publication 80-84, Jet Propulsion Laboratory, Pasadena, Calif , Dec 15, 1980

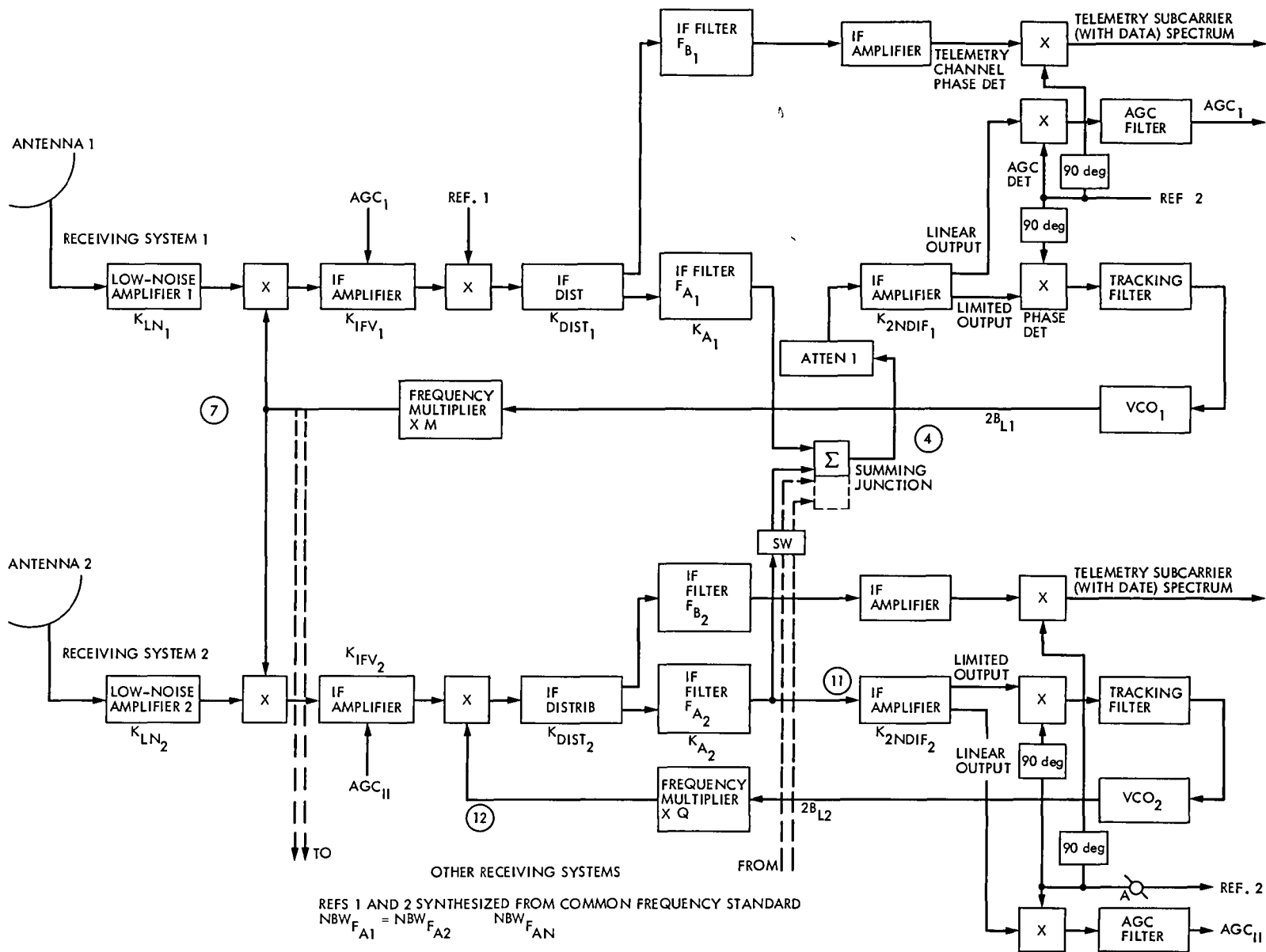


Fig 1. Radio frequency carrier arraying for high-rate telemetry reception

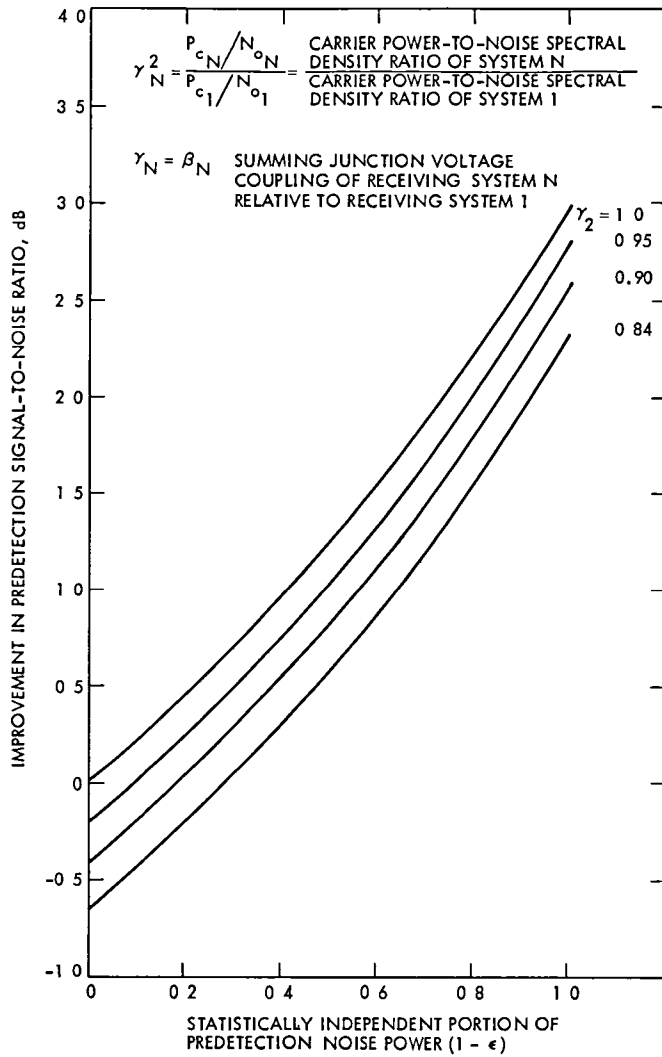


Fig 2 Effect of partial coherence in system noise temperature on predetection signal-to-noise ratio improvement, two receiving systems arrayed

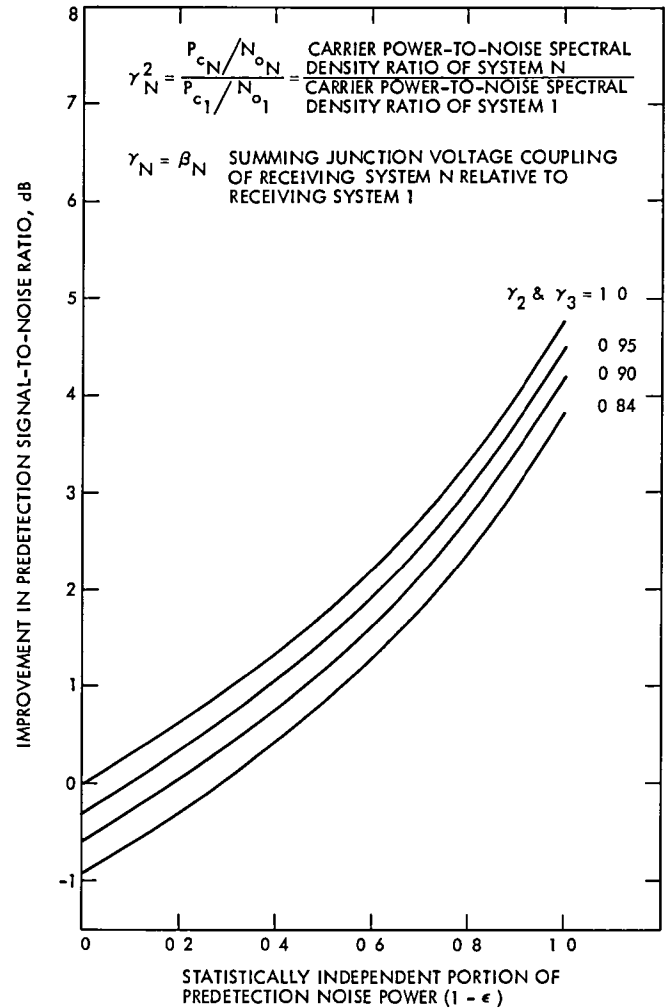


Fig. 3. Effect of partial coherence in system noise temperature on predetection signal-to-noise ratio improvement, three receiving systems arrayed

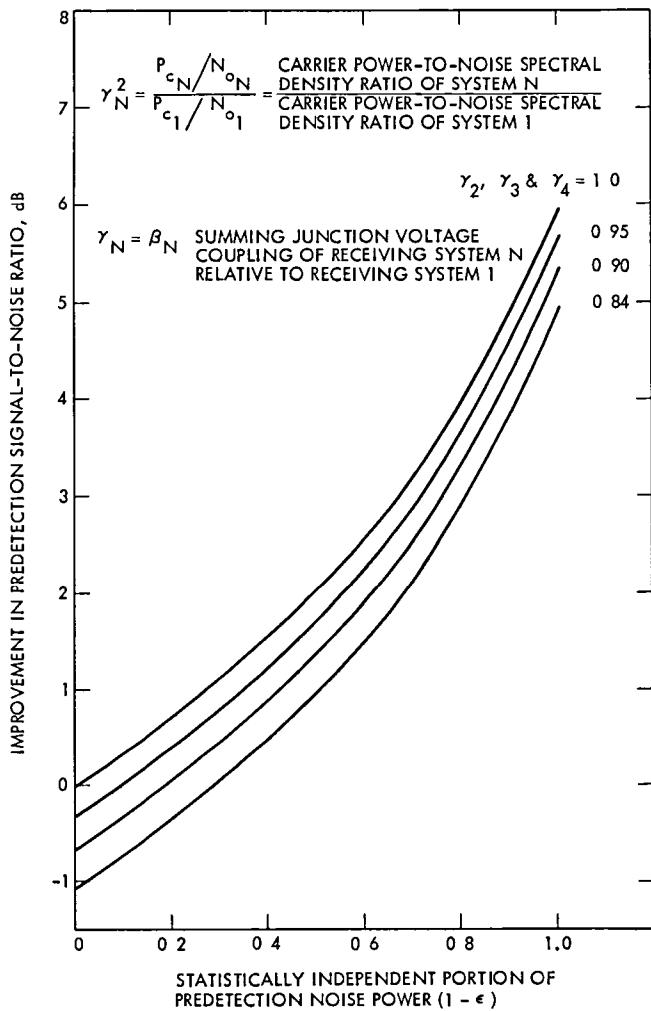


Fig. 4. Effect of partial coherence in system noise temperature on predetection signal-to-noise ratio improvement, four receiving systems arrayed

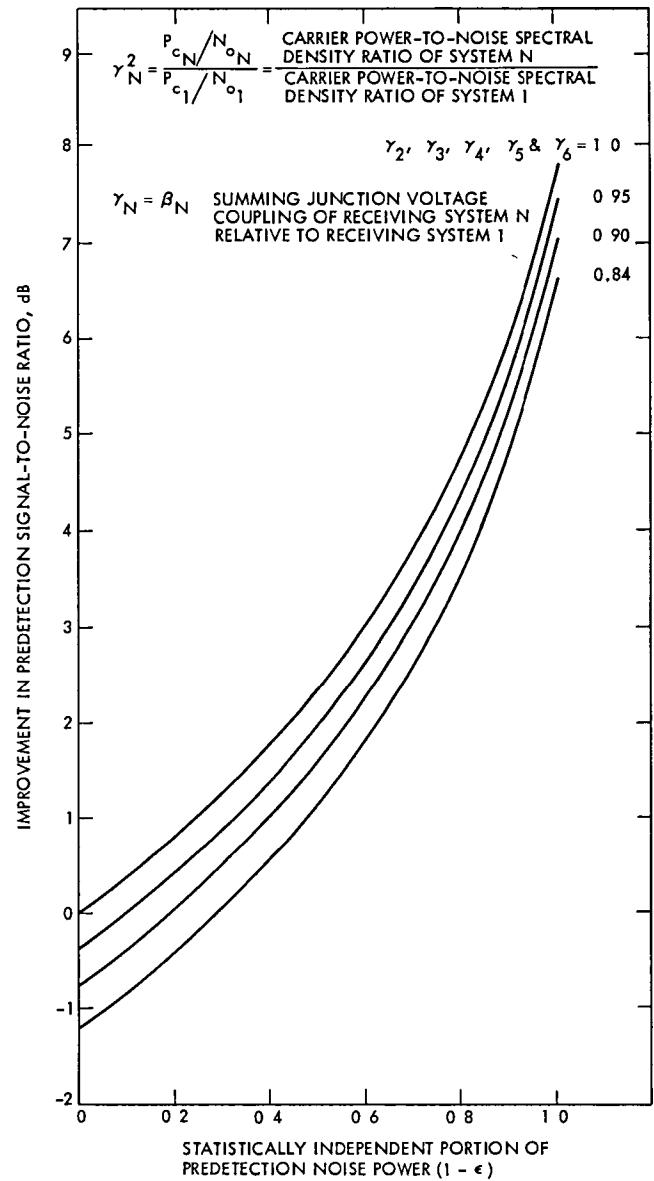


Fig 5 Effect of partial coherence in system noise temperature on predetection signal-to-noise ratio improvement, six receiving systems arrayed

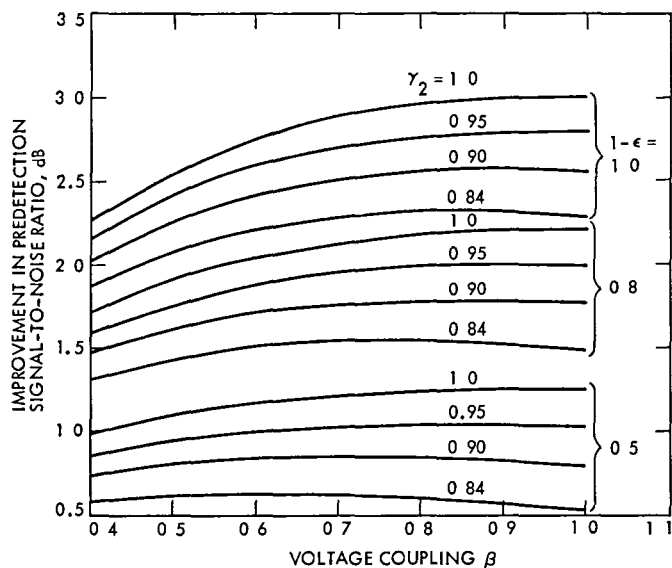


Fig. 6. Effect of summing junction voltage coupling on predetection signal-to-noise ratio improvement with partial coherence in system noise temperature, two receiving systems arrayed

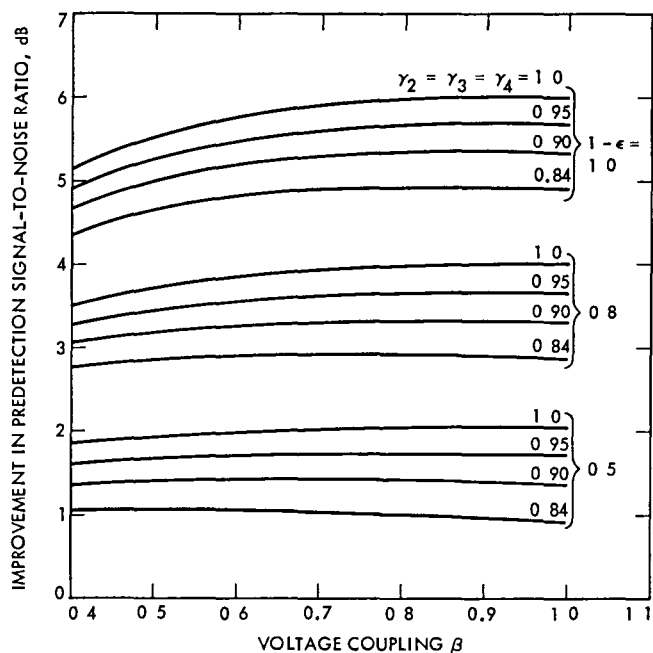


Fig. 8. Effect of summing function voltage coupling on predetection signal-to-noise ratio improvement with partial coherence in system noise temperature, four receiving systems arrayed

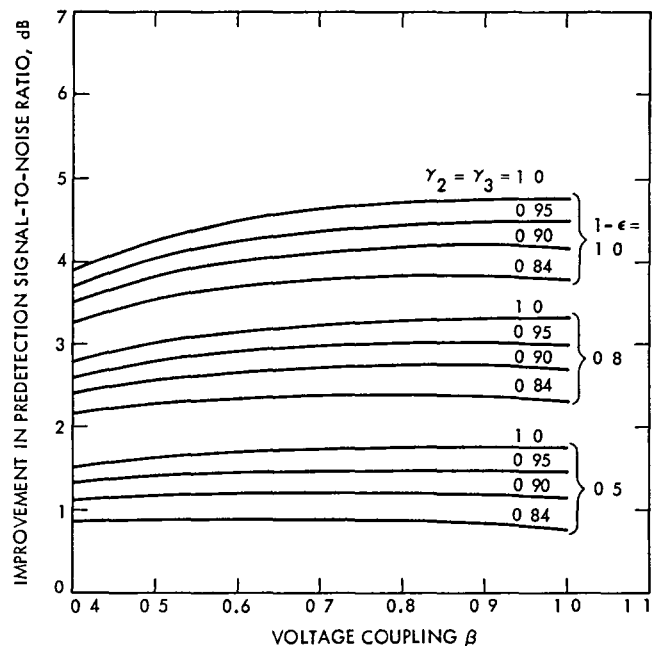


Fig. 7. Effect of summing junction voltage coupling on predetection signal-to-noise ratio improvement with partial coherence in system noise temperature, three receiving systems arrayed

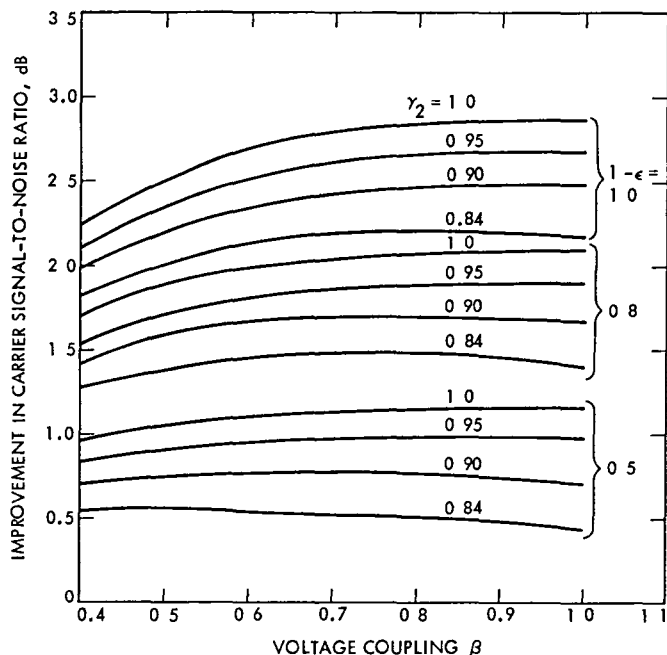


Fig. 9. Effect of summing junction voltage coupling on carrier signal-to-noise ratio improvement with partial coherence in system noise temperature, two receiving systems arrayed, high-rate telemetry configuration

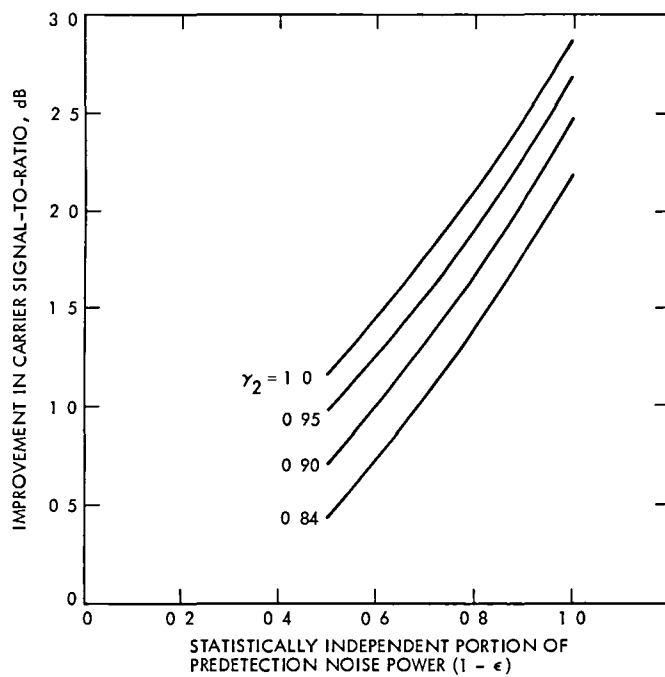


Fig 10a Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, two receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 1$

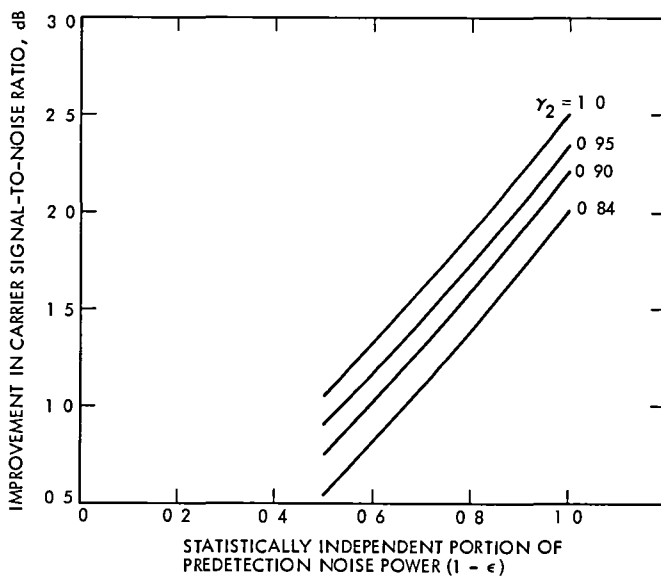


Fig. 10b. Effect of partial coherence in system noise temperature in carrier signal-to-noise ratio improvement, two receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 0.5$

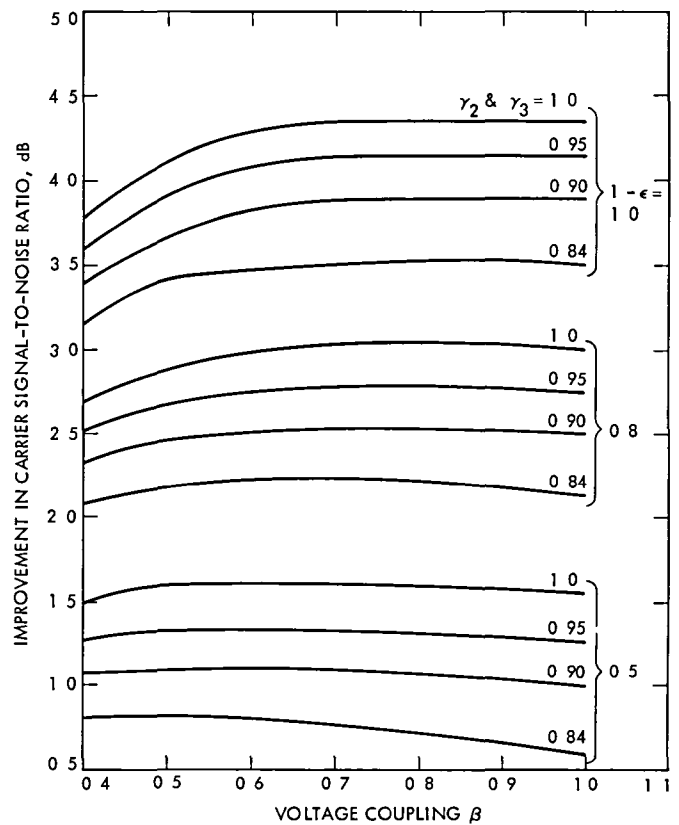


Fig. 11. Effect of summing junction voltage coupling on carrier signal-to-noise ratio improvement with partial coherence in system noise temperature, three receiving systems arrayed, high-rate telemetry configuration

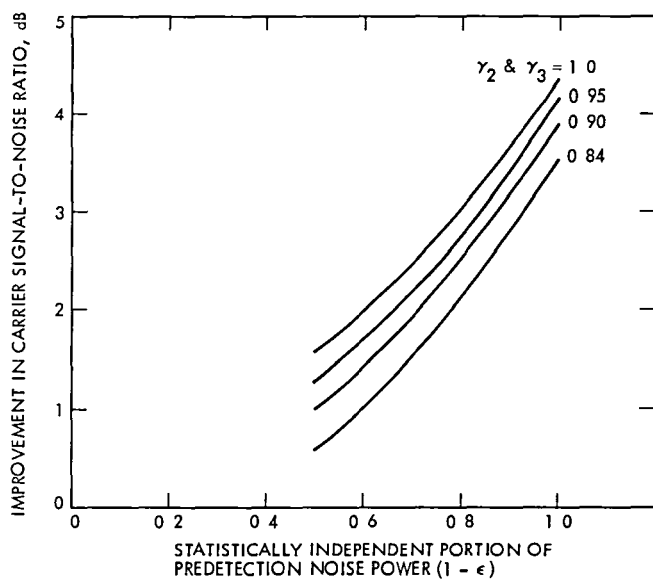


Fig 12a. Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, three receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 1$

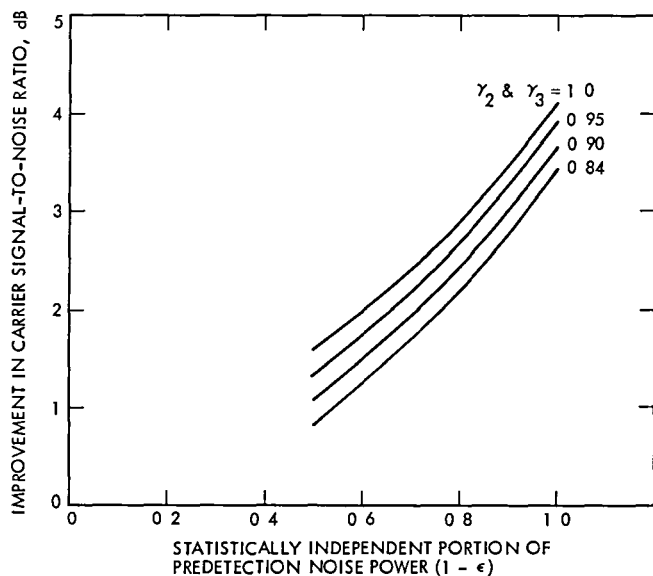


Fig. 12b. Effect of partial coherence in system noise temperature in carrier signal-to-noise ratio improvement, three receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 0.5$

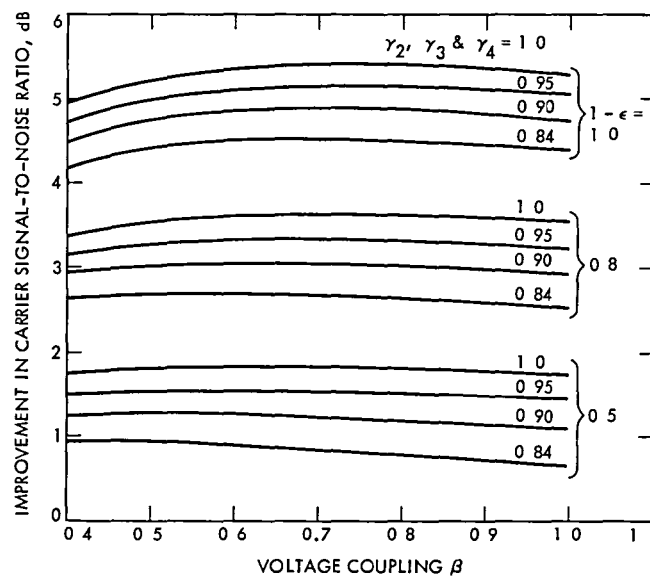


Fig 13 Effect of summing junction voltage coupling on carrier signal-to-noise ratio improvement with partial coherence in system noise temperature, four receiving systems arrayed, high-rate telemetry configuration

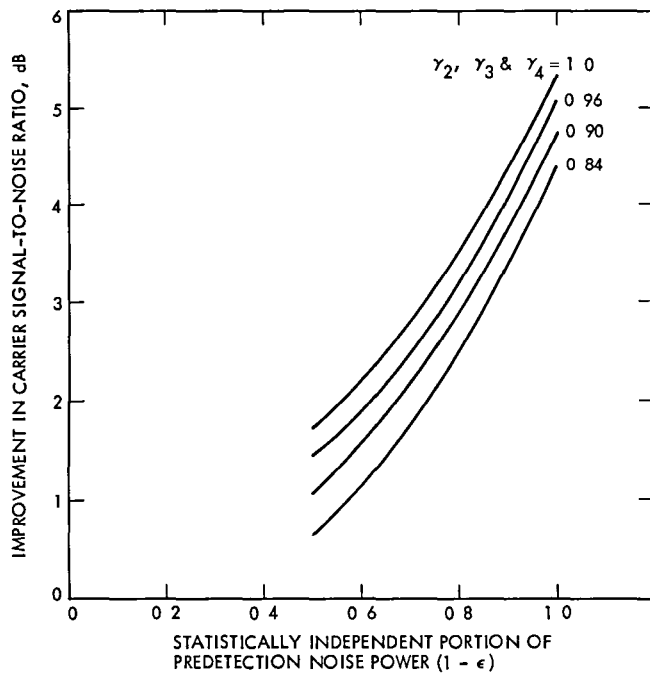


Fig 14a Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, four receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 1$

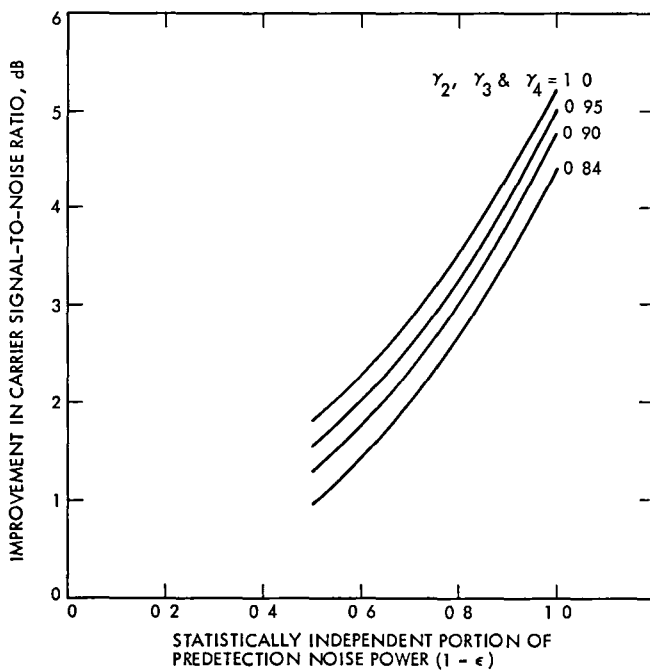


Fig 14b Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, four receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 0.5$

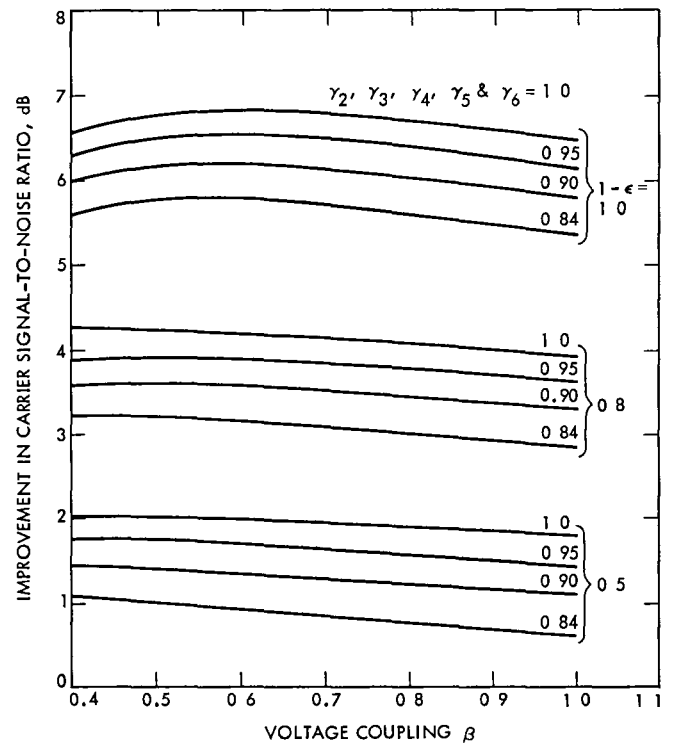


Fig 15 Effect of summing junction voltage coupling on carrier signal-to-noise ratio improvement with partial coherence in system noise temperature, six receiving systems arrayed, high-rate telemetry configuration

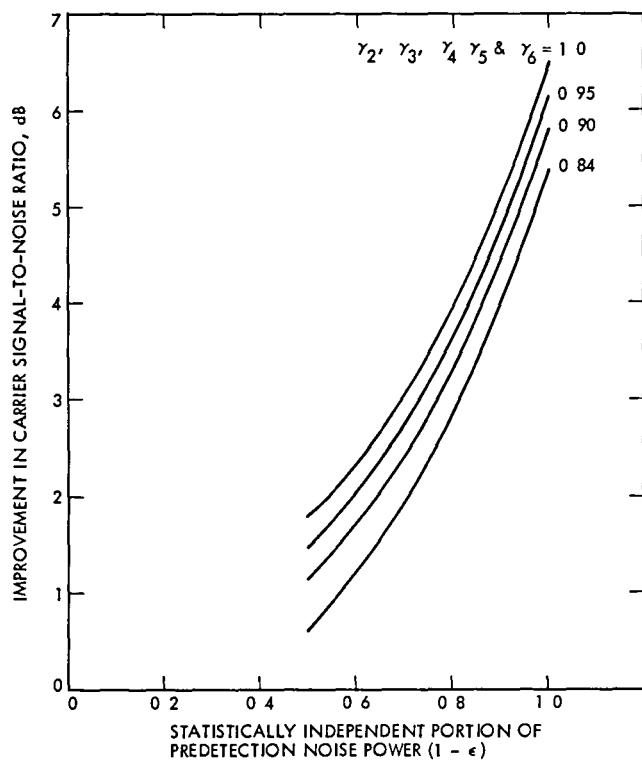


Fig. 16a Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, six receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 1$

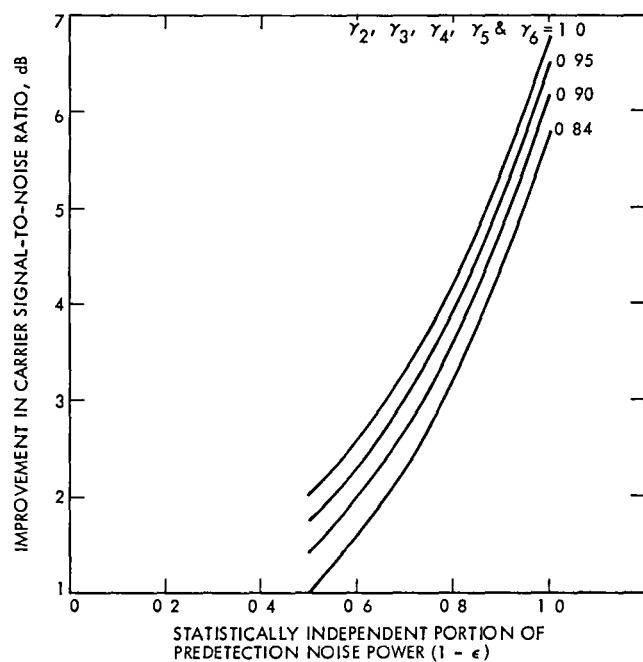


Fig. 16b Effect of partial coherence in system noise temperature on carrier signal-to-noise ratio improvement, six receiving systems arrayed, high-rate telemetry configuration, voltage coupling $\beta = 0.5$

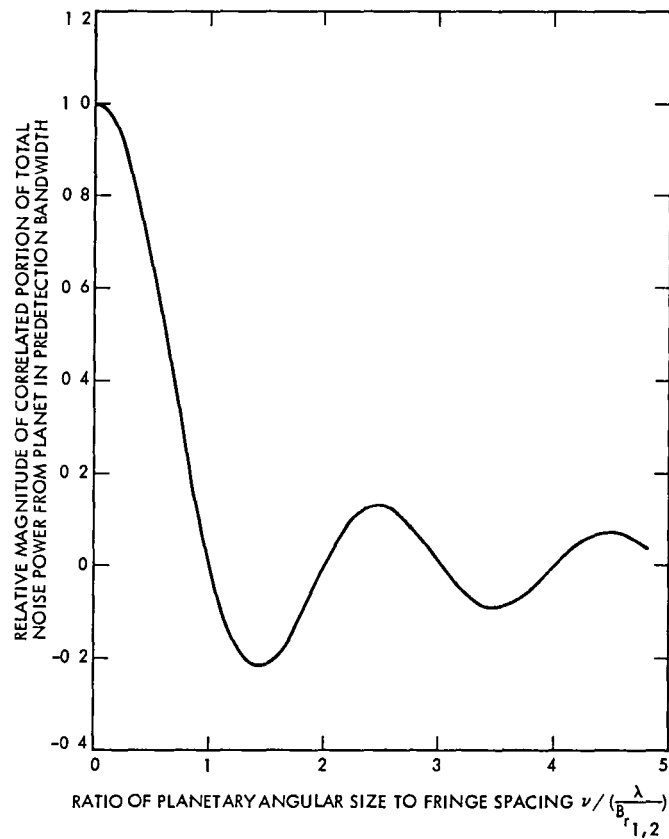


Fig. 17. Relative effect of ratio of planet angular size to fringe spacing on correlated predetection noise, two antenna array

Optimum Equipment Maintenance/Replacement Policy

Part 1: Dynamic Programming Approach

T Charng
DSN Engineering Section

This is the first part of a two-part article on the optimization of equipment maintenance and/or replacement policy. Over a given life-span of equipment, optimum policy is determined based on present costs, inflation rates, operating characteristics, future equipment developments, and other factors. A computer program utilizing the dynamic programming technique together with a numerical example is included.

Part II will present an optimum policy determined by a stochastic model such as the Markovian decision processes. In such a model, a long-term expected return/cost is optimized including the probability of equipment breakdown as functions of some performance criteria.

I. Introduction

One of the basic tasks of operating and maintenance engineering is the replacement of old machinery and obsolete tools by new and modern ones. As equipment deteriorates with use, or degrades relative to the performance of newer models or improvements in design, there comes a time for operation and maintenance engineers to make decisions regarding replacement or maintenance. When the outlay for new equipment is made, the financial loss due to the stoppage of operation and the cost of teaching new skills should be compensated for in the trade-off studies by an increase in productivity and a decrease in maintenance and operating costs.

The optimal repair and/or replacement policies under various assumptions concerning present costs, discount rates, operating characteristics, and future equipment developments are essential to the successful operation of a facility. Since decisions of this type must be made periodically, depending upon

the fundamental time period and the type of process in consideration, it is clear that this maintenance policy is a multistage decision process of the dynamic programming type.

Dynamic programming (or dynamic optimization) is a mathematical technique often used for making a sequence of interrelated decisions (Refs 1 and 2), provides a systematic procedure for determining the combination of decisions that optimizes the overall effectiveness measure. It is a general approach to many dynamic problems, but the particular equations used must be developed to fit each individual problem.

II. Computational Algorithm

The optimal equipment-replacement policies and the decision process using dynamic programming may be described as follows:

It is assumed that a machine or a piece of equipment requires a certain initial capital, generates a certain revenue, requires a certain amount of care, and can be replaced by a new machine at any time. The revenue, the maintenance cost, and the rebate on trade-in are all taken to depend on the age of the machine in known fashions. With this information, a best set of decisions, such as to KEEP, to OVERHAUL, or to REPLACE, need to be determined in order to optimize the total profit/cost over a given number of years. Thus, the objective is to optimize the total profit/cost, Y ,

$$Y = \sum_{j=1}^n F_j(i) \quad \text{for } j = 1, 2, \dots, n \quad (1)$$

with respect to decisions, D_j , where D_j can be "Keep", "Overhaul", or "Replace" for $j = 1, 2, \dots, n$ years. The function $F_j(i)$ is defined as the value at year j of the overall profit/cost from a machine that is (i) years old, where an optimal policy is employed for the remainder of process operation left.

It is assumed that the process lasts n years, and then stops. Hence,

$$F_{n+1}(i) = 0 \quad (2)$$

By the principle of optimality, the overall profit/cost of the j^{th} time period with a given decision can be obtained by

$$F_j^*(i) = \text{OPT} \{F_j^K(i), F_j^O(i), F_j^R(i)\} \quad (3)$$

Optimum function $F_j^*(i)$ can be determined by

$$\text{Keep } F_j^K(i) = R_j(i+1) - U_j(i+1) + F_{j-1}(i+1) \quad (4)$$

$$\begin{aligned} \text{Overhaul } F_j^O(i) &= R_j(k+1) - U_j(k+1) - O_j(i+1) \\ &\quad + F_{j-1}(k+1) \end{aligned} \quad (5)$$

$$\text{Replace } F_j^R(i) = R_1(1) - U_1(1) - C_j(i+1) + F_{j-1}^*(1) \quad (6)$$

where (i) is the age of the equipment at the end of the previous operation period, (k) is the equivalent age of the equipment after the overhaul, and (j) is the current operation period. The revenue R , operating/maintenance U , overhaul O , and replacement C , cost information may be tabulated or represented by analytical expressions.

III. Numerical Example

Suppose a piece of equipment (t) years old with no salvage or resale value has the annual income R in thousands of dollars, R , after all running expenses have been met, is given by

$$R(t) = \begin{cases} 25 - t^2 & \text{for } 0 < t \leq 4 \\ 0 & \text{for } t > 4 \end{cases} \quad (7)$$

Let the cost of replacement C be \$21,000, and the equipment life be 5 operating years, then the optimal future equipment replacement policy must be determined if the present equipment is 2-years old and if a decision must be made annually on the basis of a maximum total 5-year profit.

The annual profit $P_j(i)$ will be, then, the difference between income and costs

$$P_j(i) = \begin{cases} 25 - i^2 & \text{for } i \leq 4 \\ 0 & \text{for } i > 4 \end{cases} \quad (8)$$

if we keep the unit. If we replace it, the age of the equipment is $i = 0$, and

$$P_j(0) = 25 - (0)^2 - 21 = 4 \quad (9)$$

where the additional charge is the cost of replacement. Since the unit has no resale value, the net profit from a new unit is independent of the age of the unit being replaced, and we may write $P_j(0) = 4$ as the first-year profit.

We must now find the best set of decisions, KEEP (K), OVERHAUL (O), or REPLACE (R) to maximize the profit from the 5-year process life. Thus we wish to maximize

$$Y = \sum_{j=1}^5 P_j(i) \quad (10)$$

with respect to D_j , where D_j can take the value K_j or R_j for $j = 1, 2, 3, 4, 5$.

By a direct approach we could write all possible decision considerations and find the optimum policy. Thus, as in Table 1, one arbitrary combination $K-R-K-K-R$ gives a total profit of \$74,000.

This direct approach requires 2^n evaluations, or $(2)^5 = 32$ for the example, all of which need to be compared with each other to determine the largest for the 5-year period.

Since the problem is a multistage process, it can be carried out more simply by dynamic programming with a recursive function

$$F_j^*(i) = OPT_{K, R_j} \{P_j(i+1) + F_{j-1}^*(i+1)\} \quad (11)$$

For example, for a 3-year old item with 2 years to go, we will have the form

$$F_2(3) = OPT \begin{cases} P_2(3) + f_1^*(4) & \text{if keep} \\ P_2(0) + f_1^*(1) & \text{if replace} \end{cases} \quad (12)$$

The numerical value of the two possibilities are

$$F_2(3) = \begin{cases} 16 + 9 = 25 & \text{if keep} \\ 4 + 24 = 28 & \text{if replace} \end{cases} \quad (13)$$

It is therefore more profitable to replace this equipment in this situation

The complete solution using dynamic programming is given in Appendix A. For the given numerical example, there are two alternate policies. One policy is *K-K-R-K-K*, the other is *K-R-K-K-K*. Both policies result in a 5-year profit of \$86,000.

A comparison of the two policies show that more profit is returned in the first 2 years of the 5-year period with the first optimal policy than the second, and this might be preferred if the accuracy of the cost estimates might be questioned after several years of operation.

IV. Summary

A computer program utilizing the dynamic programming techniques for optimal equipment replacement policy is written in the BASIC language and is tested on a Hewlett-Packard 2647A graphics terminal. The computer program listing is given in Appendix B.

The computer program is capable of performing optimal repair and replacement policies over a given operating period under various assumptions concerning present costs, discount rates, operating characteristics, and future technology development. It provides a tool for management to decide replacement schedules and to know when and how much capital outlay will be required.

Another optimization technique including the probability of equipment break-down as a function of year, age, and other factors will be presented in Part II at a later date.

Acknowledgment

The author would like to acknowledge Dr. F. L. Lansing who made a number of helpful suggestions in the study preparation.

References

1. Bellman, R. E., Dreyfus, S. E., *Applied Dynamic Programming*, Princeton University Press, Princeton, N.J., 1962.
2. Beveridge, G. S. G., Schechter, R. S., *Optimization. Theory and Practice*, McGraw-Hill Book Company, 1970.

Table 1 Direct solution of numerical example

Age at end of previous period	Policy	Age at start of period	Profit, \$K
2	<i>K</i>	2	21
3	<i>R</i>	0	4
1	<i>K</i>	1	24
2	<i>K</i>	2	21
3	<i>R</i>	0	4
Total			74

Appendix A

Numerical Example Printout

* * * * * INPUT INFORMATION * * * * *

```

OUTPUT OPTION                      0
MACHINE STARTS AT AGE              2
AGE RETURN-BACK AFTER OVERHAUL     2
NUMBER OF TIME PERIODS             5
DISCOUNT PERCENTAGE              0
MAXIMIZATION FLAG                  1
TABULATED COST/RETURN INPUT FLAG   0

```

* * * * *

POLICY 1			
PERIOD	EQUIP AGE	DECISION	VALUE
1	2	KEEP	21
2	3	KEEP	16
3	4	REPLACE	4
4	1	KEEP	24
5	2	KEEP	21

TOTAL COST/PROFIT OVER 5 PERIODS 86

POLICY 2			
PERIOD	EQUIP AGE	DECISION	VALUE
1	2	KEEP	21
2	3	REPLACE	4
3	1	KEEP	24
4	2	KEEP	21
5	3	KEEP	16

TOTAL COST/PROFIT OVER 5 PERIODS 86

Appendix B

Computer Program Listing

```

10 REM REM
15 REM REM
20 REM EQUIPMENT REPLACEMENT REM
25 REM BY DYNAMIC PROGRAMING APPROACH REM
30 REM REM
35 REM T CHARNG 1 OCTOBER 1981 REM
40 REM REM
45 REM REM
50 DIM R(10,10),U(10,10),C(10,10),O(10,10),D(10,10,3),Opt(10,10)
55 DIM Tie(10,10),Iloc(10),P(10),F(3)
60 REM
65 INTEGER Age,Period,D,Tie,Back,Alter,P
70 REM ASSIGN PRINTOUT FILE "OUTPUT" REM
75 Ko=6
80 ASSIGN "OUTPUT" TO #Ko
85 RLM
100 REM REM
105 REM INPUT SECTION REM
110 READ Iprint ' PRINTOUT OPTION
115 READ Age ' AGE OF THE MACHINE TO BEGIN WITH
120 READ Nu ' AGE OF THE MACHINE AFTER OVERHAUL
125 READ Period ' NUMBER OF THE TIME PERIODS
130 READ Disc ' DISCOUNT PERCENTAGE
135 READ Maxi ' MAXIMIZATION FLAG
140 READ Itable ' TABULATED COST/RETURN INPUTS
145 REM PRINTOUT GENERAL INFORMATION REM
150 IF Iprint>=0 THEN GOSUB 5400
160 REM TABULATED COST/RETURN REM
165 IF Itable>0 THEN GOSUB 1000
170 REM REM
175 PRINT #Ko,LIN(?)
180 PRINT #Ko,"* * * * * "
185 REM REM
190 IF Itable<=0 THEN GOSUB 4000
195 REM REM
200 Disc=1/(1+Disc/100)
203 Ierr=0
205 REM REM
210 FOR J=1 TO Period
215 Ilast=Age+Period-J
220 FOR I=1 TO Ilast
225 Tie(1,J)=0
230 IF J<>1 THEN 260
235 F(1)=R(I+1,1)-U(1+1,1)
240 F(2)=R(Nu+1,1)-U(Nu+1,1)-O(I+1,1)
245 F(3)=R(1,1)-U(1,1)-C(I+1,1)
250 GOTO 280
255 REM
260 F(1)=R(I+1,J)-U(I+1,J)+Opt(I+1,J-1)*Disc
270 F(2)=R(Nu+1,J)-U(Nu+1,J)-O(I+1,J)+Opt(Nu+1,J-1)*Disc
275 F(3)=R(1,1)-U(1,1)-C(I+1,J)+Opt(1,J-1)*Disc
280 IF Iprint>1 THEN GOSUB 5100

```

```

290 REM
295 Opt(I,J)=F(1)
300 D(I,J,1) 1
305 FOR K=2 TO 3
310 IF Max1>0 THEN 350
315 IF Opt(I,J)<=F(K) THEN 360
320 Opt(I,J)=F(K)
325 FOR L=1 TO K-1
330 D(I,J,L)=0
335 NEXT L
340 D(I,J,K)=K
345 GOTO 375
350 IF Opt(I,J)>=F(K) THEN 360
355 GOTO 320
360 IF Opt(I,J)<>F(K) THEN 375
365 D(I,J,K)=K
370 Tie(I,J)=Tie(I,J)+1
375 NEXT K
380 NEXT I
385 PRINT #Ko
390 NEXT J
400 REM
410 Alter=1
415 IF lprint>1 THEN GOSUB 5020
420 Back=0
425 I1=0
430 I=Age
435 J=Period
440 REM
445 I1=I1+1
450 lloc(I1)=I
455 V(J)=Opt(I,J)
460 FOR K=1 TO 3
465 IF D(I,J,K)<=0 THEN 535
470 P(J)=D(I,J,K)
475 IF D(I,J,K)<>1 THEN 495
480 IF lprint>1 THEN PRINT #Ko,TAB(20),J,I,"KEEP",Opt(I,J)
485 I=I+1
490 GOTO 560
495 IF D(I,J,K)<>2 THEN 515
500 IF lprint>1 THEN PRINT #Ko,TAB(20),J,I,"OVERHAUL",Opt(I,J)
505 I=Nu
510 GOTO 560
515 IF D(I,J,K)<>3 THEN 540
520 IF lprint>1 THEN PRINT #Ko,TAB(20),J,I,"REPLACE",Opt(I,J)
525 I=1
530 GOTO 560
535 NEXT K
540 IF lerr>0 THEN 65000
541 lerr=1
542 lprint=2
545 PRINT #Ko,TAB(5),"$$$$$$$$$  SOMETHING IS WRONG AT PERIOD",J,
546 PRINT #Ko,"AGE",I,"D(",I,",",J,",",K,")=",D(I,J,K)
547 PRINT #Ko,LIN(1),"EXAMINE THE RESULTS CAREFULLY"
548 PRINT #Ko
550 GOTO 220

```



```

560 REM REM
565 I1=Iloc(I1)
570 IF Tie(I1,J)<=0 THEN 590
575 Back=1
580 Tie(I1,J)=Tie(I1,J)-1
585 D(I1,J,K)=0
590 J=J-1
595 IF J>=1 THEN 440
600 REM REM
605 GOSUB 5000
610 FOR J=Period TO 1 STEP -1
615 Jp=Period-J+1
620 I1=Iloc(Jp)
625 Vp=V(J)
630 IF J>1 THEN Vp=Vp-V(J-1)
635 IF P(J)=1 THEN PRINT #Ko,TAB(20),Jp,I1,"KLEP",Vp
640 IF P(J)=2 THEN PRINT #Ko,TAB(20),Jp,I1,"OVERHAUL",Vp
645 IF P(J)=3 THEN PRINT #Ko,TAB(20),Jp,I1,"REPLACE",Vp
647 NEXT J
655 PRINT #Ko,TAB(20),"-----"
660 PRINT #Ko,LIN(1),TAB(20),"TOTAL COST/PROFIT OVER",Period,"PERIODS",
665 PRINT #Ko,TAB(10),Opt(Age,Period)
670 PRINT #Ko
680 REM REM
685 Alter=Alter+1
690 IF Back=1 THEN 415
695 GOTO 65000
1000 REM SUB FOR TABULATED COST/RETURN INPUTS REM
1010 FOR I=1 TO Period+2
1020 FOR J=1 TO Period
1030 READ R(I,J) I REVENUE
1040 NEXT J
1050 NEXT I
1060 REM
1070 FOR I=1 TO Period+2
1080 FOR J=1 TO Period
1090 READ U(I,J) I OPERATION/MAINTENANCE
1100 NEXT J
1110 NEXT I
1120 REM
1130 FOR I=1 TO Period+2
1140 FOR J=1 TO Period
1150 READ C(I,J) I REPLACRMENT
1160 NEXT J
1170 NEXT I
1180 REM
1190 FOR I=1 TO Period+2
1200 FOR J=1 TO Period
1210 READ O(I,J) I OVERHAUL
1220 NEXT J
1230 NEXT I
1240 RETURN

```

```

2000 REM                                INPUT DATA BLOCK                                REM
2010 DATA 0                            I PRINT(OUT OPTION                                REM
2020 DATA 2                            I AGE OF THE MACHINE TO BEGIN WITH
2030 DATA 2                            I AGE OF THE MACHINE AFTER OVERHAUL
2040 DATA 5                            I NUMBER OF THE TIME PERIODS
2050 DATA 0                            I DISCOUNT PERCENTAGE
2060 DATA 1                            I MAXIMIZATION FLAG
2070 DATA 0                            I TABULATED COST/RETURN INPUTS
2100 REM                                REVENUE MATRIX R(I,J)                                REM
2110 DATA 155,150,140,135,125
2120 DATA 140,135,125,120,115
2130 DATA 125,110,110,115,100
2140 DATA 105,105,110, 90, 95
2150 DATA 100,105, 80, 90, 70
2160 DATA 100, 70, 80, 65, 70
2170 DATA 60, 70, 65, 70, 60
2200 REM                                MAINTENANCE/OPERATION COST U(I,J)                                REM
2210 DATA 5, 5, 5, 10, 10
2220 DATA 10, 10, 10, 10, 10
2230 DATA 10, 10, 10, 15, 20
2240 DATA 10, 15, 15, 20, 20
2250 DATA 15, 20, 25, 25, 25
2260 DATA 20, 25, 25, 30, 30
2270 DATA 30, 30, 30, 35, 45
2300 REM                                REPLACEMENT COST C(I,J)                                REM
2310 DATA 220,220,220,210,210
2320 DATA 225,230,220,220,215
2330 DATA 240,230,230,220,220
2340 DATA 240,240,225,225,250
2350 DATA 250,230,230,255,255
2360 DATA 235,235,260,260,260
2370 DATA 240,265,265,265,270
2400 REM                                OVERHAUL COST O(I,J)                                REM
2410 DATA 55, 55, 55, 60, 60
2420 DATA 60, 60, 60, 60, 60
2430 DATA 60, 60, 60, 65, 70
2440 DATA 60, 65, 65, 70, 70
2450 DATA 65, 70, 75, 75, 75
2460 DATA 60, 65, 65, 80, 80
2470 DATA 80, 80, 80, 85, 95
3000 REM                                END OF DATA BLOCK                                REM
4000 REM                                REM
4020 FOR J=1 TO Period
4030 R(I,J)=25
4040 C(I,J)=21
4050 U(I,J)=0
4060 O(I,J)=10
4070 FOR I=2 TO Period+Age
4080 Im1=I-1
4090 R(I,J)=R(I,J)-Im1*Im1
4100 IF Im1>4 THEN R(I,J)=0
4110 C(I,J)=C(I,J)
4120 O(I,J)=O(I,J)
4130 U(I,J)=U(I,J)
4140 NEXT I
4150 NEXT J
4160 RETURN

```

```

5000 REM REM
5002 PRINT #K0,LIN(?)
5004 PRINT #K0,TAB(20),"----- POLICY",Alter,"-----"
5006 PRINT #K0,TAB(20),"PERIOD","EQUIP AGE","DECISION","VALUE"
5008 PRINT #K0,TAB(20),"-----","-----","-----","-----"
5010 RETURN
5020 REM REM
5022 PRINT #K0,LIN(2)
5024 PRINT #K0,TAB(20),"----- POLICY",Alter,"-----"
5026 PRINT #K0,TAB(20),"PERIOD","EQUIP AGE","DECISION","CUMULATED VALUE"
5028 PRINT #K0,TAB(20),"-----","-----","-----","-----"
5030 RETURN
5100 REM REM
5120 PRINT #K0,TAB(5),"PERIOD=",J,"EQUIP AGE=",I,
5130 PRINT #K0,TAB(5),"KEEP=",R(I+1,J),"-",U(I+1,J),
5140 PRINT #K0,"+",Opt(I+1,J-1),"*",Disc,"=",F(1)
5150 PRINT #K0,TAB(33),"RENU=",R(Nu+1,J),"-",U(Nu+1,J),"-",O(I+1,J),
5160 PRINT #K0,"+",Opt(I+1,J-1),"*",Disc,"=",F(2)
5170 PRINT #K0,TAB(33),"REPL=",R(1,J),"-",U(1,J),"-",C(I+1,J),
5180 PRINT #K0,"+",Opt(1,J-1),"*",Disc,"=",F(3)
5190 RETURN
5300 REM REM
5320 PRINT #K0,LIN(20),TAB(5)," J "," I ","D(I,J,K)","OPT(I,J)","TIE(I,J)"
5325 PRINT #K0,TAB(5),"---","---","-----","-----","-----"
5330 PRINT #K0
5335 FOR J=1 TO Period
5340 FOR I=1 TO Period+Age-J
5345 PRINT #K0,TAB(5),J,I,D(I,J,1),D(I,J,2),D(1,J,3),Opt(I,J),Tie(I,J)
5350 NEXT I
5355 PRINT #K0
5360 NEXT J
5365 RETURN
5400 REM REM
5410 PRINT #K0,LIN(2)
5415 PRINT #K0,"* * * * * INPUT INFORMATION * * * * *"
5420 PRINT #K0,LIN(?)
5425 PRINT #K0,TAB(20),"OUTPUT OPTION",Iprint
5430 PRINT #K0,TAB(20),"MACHINE STARTS AT AGE",Age
5435 PRINT #K0,TAB(20),"AGE RETURN-BACK AFTER OVERHAUL",Nu
5440 PRINT #K0,TAB(20),"NUMBER OF TIME PERIODS",Period
5445 PRINT #K0,TAB(20),"DISCOUNT PERCENTAGE",Disc
5450 PRINT #K0,TAB(20),"MAXIMIZATION FLAG",Max1
5455 PRINT #K0,TAB(20),"TABULATED COST/RETURN INPUT FLAG",Itable
5460 RETURN
5500 REM REM
65000 REM
65010 PRINT #K0,LIN(5)
65020 PRINT #K0,TAB(25),"* * * * * * * * * * * * * * * * *"
65030 PRINT #K0,TAB(25),"* * * * * * * * * * * * * * * * *"
65040 PRINT #K0,TAB(25),"* * * * * * * * * * * * * * * * *"
65050 PRINT #K0,TAB(25),"* * * * * * * * * * * * * * * * *"
65060 PRINT #K0,TAB(25),"* * * * * * * * * * * * * * * * *"
65070 REM
65080 COMMAND "M F H HP-IB#1"
65090 REM
65100 END

```

System Reliability Analysis with Dynamic Programming

T Charng
DSN Engineering Section

System reliability analysis using the "Mission Success Diagram" method is presented in this article. In addition to the basic building blocks and reliability prediction equations, a BASIC computer program is written and tested with certain system configurations. This computer program can be modified for various system structures.

I. Introduction

Reliability, $R(t)$, is the probability that a device performs adequately over time interval $(0, t)$. In general, it is assumed that unless repair or replacement occurs, adequate performance at time t implies adequate performance during the interval $(0, t)$. The device under consideration may be an entire system, a subsystem containing one or more components, or a component. In principle, it is possible to break down the system into black boxes, with each black box being in one of two states: good or bad. Mathematical models of the system can then be abstracted from the physical processes and the theory of combinatorial probability utilized to predict the reliability of the system. The black boxes may be independent of, or dependent upon, each other.

When it is proposed to design a system to perform a complex and demanding task, it is assumed that the required investment will be justified according to the perfection by which the task is performed or by the large number of times which the system can do the job. This assumption cannot be justified when a system fails to perform upon demand or fails to perform repeatedly. Thus, it is not enough to show that a chasm can be spanned by a bridge; the bridge must continue to span the chasm for a long time to come while carrying a useful load.

Reliability is a problem at all levels. Reliability engineering (Refs 1, 2) is concerned with the time degradation of materials, physical and electronic measurements, equipment design, process and system analysis, and synthesis.

To be of value, a prediction must be timely. The earlier a prediction has to be made about the unknown nature of a future event, the more difficult it is to make a meaningful prediction.

It can be seen that the reliability of a device is known with certainty after it has been used in the field until it is worn out and its failure history has been faithfully recorded. Thus a fundamental limitation on reliability prediction is the ability to accumulate data of known validity for the new application. Another fundamental limitation is the complexity of prediction techniques. Very simple techniques omit a great deal of distinguishing detail and the prediction suffers inaccuracy. More detailed techniques can become so cumbersome that the prediction becomes costly and impractical.

II. Structural Function of a System

Suppose a system can be divided into n components. The performance of each component can be denoted by a random

variable X_i , which takes on the value $X_i = 1$ if the component performs satisfactorily for the desired time and $X_i = 0$ if the component fails during this time. In general then, X_i is a binary random variable defined by

$$X_i = \begin{cases} 1, & \text{if component } i \text{ performs satisfactorily during time } (0, t) \\ 0, & \text{if component } i \text{ fails during time } (0, t) \end{cases} \quad (1)$$

The performance of the system is measured by the binary random variable $\phi(X_1, X_2, \dots, X_n)$, where

$$\phi(X_1, X_2, \dots, X_n) = \begin{cases} 1, & \text{if system performs satisfactorily during time } (0, t) \\ 0, & \text{if system fails during time } (0, t) \end{cases} \quad (2)$$

The function ϕ is called the "structure function of the system" and is just a function of the n -component random variables

Since the performance of each component in the system takes on the value 1 or 0, then the function ϕ is defined over 2^n points, with each point resulting in a 1 if the system performs satisfactorily and a zero if the system fails.

There are several important structure functions to consider, depending upon how the components are assembled. These are discussed in detail in the following

III. Reliability of Series Systems

The series system is the simplest and most common of all configurations. For a series system, the system fails if any component of the system fails, i.e., it performs satisfactorily if and only if all the components perform satisfactorily. The structure function of a system containing n components is a binary random variable that takes on the value 1 or 0. Furthermore, the reliability of this system can be expressed as

$$R = P\{\phi(X_1, X_2, \dots, X_n) = 1\} \quad (3)$$

where P is the probability of success. When the usual times for conditional probability are employed,

$$R = P\{X_1 = 1\} P\{X_2 = 1 | X_1 = 1\} P\{X_3 = 1 | X_1 = 1, X_2 = 1\} \dots P\{X_n = 1 | X_1 = 1, X_2 = 1, \dots, X_{n-1} = 1\} \quad (4)$$

where $P\{X_2 = 1 | X_1 = 1\}$ is the probability that component 2 will perform successfully, given that component 1 performs

successfully. The performances of these components are then dependent, and the evaluation of the conditional probability is extremely difficult.

If, on the other hand, the performance characteristics of these components do not interact, then the components can be said to be independent. The expression for the reliability then simplifies and becomes

$$R = P\{X_1 = 1\} P\{X_2 = 1\} \dots P\{X_n = 1\} \quad (5)$$

When components of a series system are assumed to be independent, it should be noted that the reliability is a function of the probability distribution of the X_i . This phenomenon is true for any system structure.

Unless otherwise specified it will be assumed throughout the remainder of this report that the component performances are independent. Hence, the probability distribution of the binary random variables X_i can be expressed as

$$P\{X_i = 1\} = p_i \quad (6)$$

and

$$P\{X_i = 0\} = 1 - p_i \quad (7)$$

Thus for systems composed of independent components, the reliability becomes a function of the p_i , that is,

$$R = R(p_1, p_2, \dots, p_n) = p_1 p_2 \dots p_n \quad (8)$$

IV. Reliability of Parallel Systems

A parallel system of n components is defined to be a system that fails if all components fail, or alternatively, the system performs satisfactorily if at least one of the n components performs satisfactorily (with all n components operating simultaneously). This property of parallel systems is often called "redundancy", there are alternative components, existing within the system, to keep the system operating successfully in case of failure of one or more components. The structure function for a parallel system is given by

$$\phi(X_1, X_2, \dots, X_n) = 1 - (1 - X_1)(1 - X_2) \dots (1 - X_n) \quad (9)$$

The structure function takes on the value 1 if at least one of the X_i equals 1.

As previously indicated, for systems composed of independent components the reliability is given by

$$R(p_1, p_2, \dots, p_n) = 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n) \quad (10)$$

V. System Reliability Modeling

Building blocks is a method making use of the equations developed for redundancy to handle series, parallel combinations of equipments. For non-series parallel, series-parallel or complex configurations, use or repeated use of the following equation is required.

$$R = R(\text{if } X \text{ is good}) R_x + R(\text{if } X \text{ is bad}) Q_x \quad (11)$$

where R is reliability of the system, R_x is reliability of component X , and Q_x is unreliability of component $X = 1 - R_x$

In order to construct a reliability model, procedures for one method are listed as follows

- (1) Define what is required for mission success and translate this into a "mission success diagram"
- (2) Write the probability of success equation for the system
- (3) Calculate the probability of success of the equipment in the system by utilizing one of the various reliability prediction techniques
- (4) The probability numbers derived in Step 3 are inserted in the formula derived in Step 2
- (5) A probability of success curve versus time can be plotted by taking several values of time for mission time, and evaluating the probability of system success by the above procedure for the several values of time chosen
- (6) Additional steps in the analysis will depend upon the decisions that the analysis is intended to be optimized

VI. Building Blocks of the System Reliability

Equations for basic building blocks of the probability of success are presented as follows for various "mission success diagrams". Each formula shown can be used as a building block to evaluate a more complex mission success diagram

If there is only one component in the system, then the mission success diagram is as shown in Fig 1(a). The reliability

for the system is obviously the probability of success of equipment A, or

$$R = p_A \quad (12)$$

For two components in series, the mission success diagram is shown in Fig 1(b), and the system reliability is

$$R = R(\text{if } A \text{ is good}) p_A + R(\text{if } A \text{ is bad}) (1 - p_A) \quad (13)$$

or

$$R = (p_B) (p_A) + (0) (1 - p_A) = p_A p_B \quad (14)$$

For two components in parallel, the mission success diagram is shown in Fig 1(c)

The reliability of the system is

$$R = R(\text{if } A \text{ is good}) p_A + R(\text{if } A \text{ is bad}) (1 - p_A) \quad (13)$$

or

$$R = (1) (p_A) + (p_B) (1 - p_A) = p_A + p_B - p_A p_B \quad (15)$$

For a complex system with the mission success diagram as in Fig 2(a)

The system reliability can be expressed as

$$R = R(\text{if } A \text{ is good}) p_A + R(\text{if } A \text{ is bad}) (1 - p_A) \quad (13)$$

When A is good, the system may be viewed as in Fig 2(b), and

$$R(\text{if } A \text{ is good}) = p_{C1} + p_{C2} - p_{C1} p_{C2} \quad (16)$$

If A is bad, then the system may be viewed as in Fig 2(c), and

$$R(\text{if } A \text{ is bad}) = p_{B1} p_{C1} + p_{B2} p_{C2} - p_{B1} p_{C1} p_{B2} p_{C2} \quad (17)$$

Thus the system reliability is

$$R = (p_{C1} + p_{C2} - p_{C1} p_{C2}) p_A + (p_{B1} p_{C1} + p_{B2} p_{C2}) (1 - p_A) \quad (18)$$

VII. Computer Program

A computer program for the system reliability analysis written in the BASIC language utilizes dynamic programming technique (Refs 3-5). In contrast to other optimization techniques (such as linear programming), a standard mathematical formulation does not exist. Dynamic programming starts with a small portion of the problem and finds the optimal solution for this smaller problem. It then gradually enlarges the problem, finding the current optimal solution from the previous one, until the original problem is solved in its entirety. Consequently, dynamic programming is a general strategy for optimization rather than a specific set of rules; the particular set of equations used must be developed to fit each problem.

The system reliability can be improved by providing redundancy in one or more of the components as mentioned before. If the probability of that component (N) functioning is P , then the probability that it will function with X units in parallel will be,

$$R(N, X) = P(N, X) \quad (19)$$

The probability that all components in the system will function will be the product of these probabilities for all components. The installation of these components will usually involve additional cost. The number of parallel units may be constrained by the size and weight of the system in addition to budgetary constraints.

In dynamic programming terminology, the STATE S of the system is the number of dollars available to a particular component to improve its reliability. The components will correspond to the STAGE N . The DECISION X will be the number of units of a component to be installed in parallel.

The recursive relation for this problem is

$$f_N^* = OPT \{f_N(N, S, X)\} = R(N, X) f_{N-1}^* [S - M(N, X)] \quad (20)$$

where f^* is the optimum value determined by the S , N , and X .

The probabilities P and the cost M for each value of X are in arrays of P and M of the computer program.

In the case where the probability of each component with parallel installation is not determined experimentally, a similar recursive relation uses theoretical estimates for the probability of success instead of using tabulated probability values. The modified recursive relation uses theoretical estimates for the probability of success instead of tabulated probability values.

The modified recursive function is

$$f_N^* = OPT \{f_N(N, S, X)\} \\ = [1 - R(N, 1)] f_{N-1}^* [S - M(N, X)] \quad (21)$$

where

$$R(N, 1) = 1 - P(N, 1)^X \quad (22)$$

The computer program listing is given in Appendix A. The description of the technique may be illustrated with the following two examples.

VIII. Numerical Examples

First let us consider the reliability problem with a system composed of four components as shown in Fig 3. Each component has the probability of functioning with 1, 2, and 3 parallel units as given in Table 1. The cost of installing 1, 2, or 3 units in each of the components is tabulated in Table 2. The constraint is that no more than \$45K can be spent on the system. The question is how many parallel units in each component should be installed in order to maximize the system reliability?

The input information and computer printout is given in Appendix B. The solution of the problem is

Total cost (in \$1000)	45
System reliability	0.22
Redundancy of components	2 units of component 1 2 units of component 2 2 units of component 3 1 unit of component 4

The second example, as shown in Fig 4, is similar to the first example except that the probability of functioning for parallel components is not known. A theoretical estimate for the probability of success is expressed by Eq (22). The recursive relation is given by Eq (21).

With the input information given in Appendix C, the solution of a maximum of four parallel units for each component allowed is that

Total cost (in \$1000)	45
System reliability	0.28
Redundancy of components	2 units of component 1 2 units of component 2 1 unit of component 3 2 units of component 4

IX. Summary

Building blocks and reliability equations based on the "mission success diagram" method is presented. Two numerical examples were solved with a Hewlett-Packard 2647A graphics terminal. The computer program is proved capable of analyzing system reliability by providing redundancy in

one or more of the components subject to size and weight of the system in addition to budgetary constraints.

This computer program provides a main framework based on the dynamic programming optimization technique. More generalized recursive relations or additional recursive relations may be required for different system configurations.

Acknowledgment

The author would like to acknowledge Dr. F. L. Lansing, who made a number of helpful suggestions in the study preparation.

References

1. Hillies, F. S., and Lieberman, G. J., *Introduction to Operations Research*, 3rd Edition, Holden-Day, Inc., San Francisco, CA, 1980.
2. Military Standardization Handbook, *Reliability Prediction of Electronic Equipment*, U.S. Dept. of Defense, 1978.
3. Bellman, R. E., and Dreyfus, S. E., *Applied Dynamic Programming*, Princeton University Press, Princeton, N.J., 1962.
4. Kuester, J. L., and Mize, J. H., *Optimization Techniques with Fortran*, McGraw-Hill Book Co., N.Y., 1973.
5. Rider, E., *General Computer Solution of Dynamic Programming Problems with Integer Restrictions*, M.S. thesis, Arizona State University, 1971.

Table 1 Probability of functioning for example 1

Component	Number of parallel units		
	1	2	3
1	0.60	0.75	0.85
2	0.40	0.65	0.80
3	0.70	0.90	0.95
4	0.50	0.60	0.80

Table 2 Costs of redundancy for example 1, in \$1000

Component	Number of parallel units		
	1	2	3
1	\$ 6	\$11	\$15
2	10	16	22
3	5	10	14
4	8	13	17

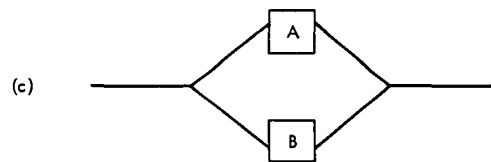
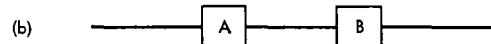
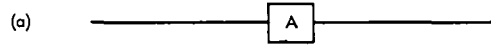


Fig 1 Basic building blocks for mission success diagram

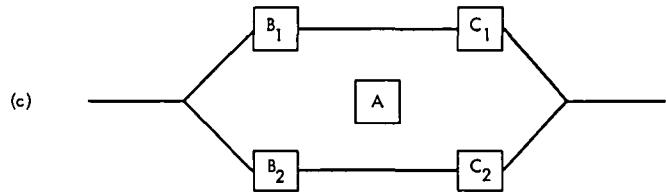
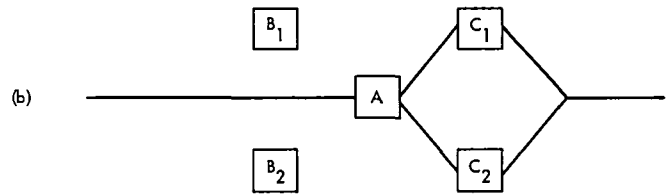
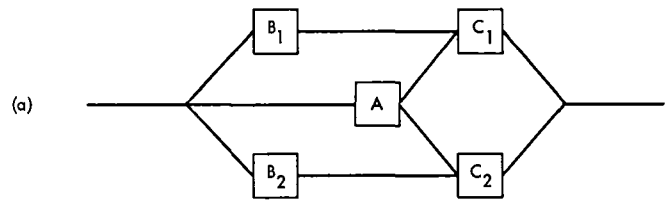


Fig 2 A complex mission success diagram

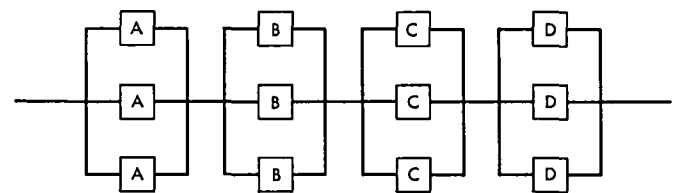


Fig 3 Mission success diagram for example 1

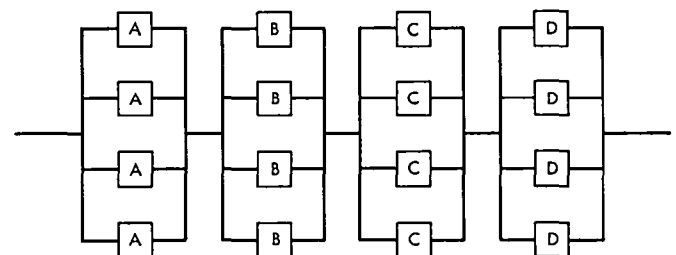


Fig 4 Mission success diagram for example 2

Appendix A

Program Listing

```

50 REM REM
51 REM REM
52 REM GENERALIZED DYNAMIC PROGRAMMING PACKAGE REM
54 REM POSITIVE VALUED INTEGER SOLUTIONS ONLY REM
55 REM REM
56 REM T CHARNG 1 OCTOBER 1981 REM
57 REM REM
70 REM REM
71 REM DIM Fsbest(Maxsize),Idbest(Maxsize),Fsopt(Maxsize) REM
72 REM DIM Kdec(Nstage),Kties(Nstage),Mtb(Nstage) REM
73 REM DIM Lps(Nstage),Mps(Nps),Lpd(Nstage),Mpd(Npd) REM
74 REM DIM P(Nstage,Ncolsp),C(Nrowsc,Ncolsc),M(Nstage,Ncolsm) REM
80 REM REM
81 DIM Fsbest(100),Idbest(100),Fsopt(100)
82 DIM Kdec(10),Kties(10),Mtb(10)
83 DIM Lps(10),Mps(20),Lpd(10),Mpd(20)
84 DIM P(10,10),C(10,10),M(10,10)
85 DIM Comment$[170]
90 REM REM
91 INTEGER Hs,Hd,Funct
99 REM DATA BLOCK REM
100 REM ENTER DATA FUNCT, MIN, IOPT3, IOPT2, IOPT1
150 REM ENTER DATA NSTAGE, LOWD, MAXD, LOWS, MAXS, INC
200 REM ENTER DATA NCOLSP, NCOLSM, NPD, NPS, NROWSC, NCOLSC
250 REM ENTER DATA ( P(I,J), J=1,NCOLSP), I=1,NSTAGE )
350 REM ENTER DATA ( M(I,J), J=1,NCOLSM), I=1,NSTAGE )
400 REM ENTER DATA (MPD(I), I=1,NPD)
450 REM ENTER DATA (MPS(I), I=1,NPS)
500 REM ENTER DATA (MTB(I), I=1,NSTAGE)
550 REM ENTER DATA (FSOPT(I), I=1,NNS)
600 REM ENTER DATA (IDBEST(I), I=1,NNS)
900 REM ENTER COMMENTS
999 REM END OF DATA BLOCK REM
1000 REM
1010 REM OPEN READ/PRINT FILES REM
1012 K1=5
1014 REM ASSIGN "INPUT" TO #K1
1015 K0=6
1020 ASSIGN "OUTPUT" TO #K0
1050 REM PROGRAM STARTS REM
1125 Maxsize=100
1140 READ Funct,Min,Iopt3,Iopt2,Iopt1
1145 Bigm=-9E+37
1150 IF Min>0 THEN Bigm=-Bigm
1160 IF Funct=4 THEN Iopt3=1
1170 IF Funct=5 THEN Iopt3=1
1172 PRINT #K0,LIN(1),TAB(5),"PROBLEM TYPE",Funct
1173 PRINT #K0,TAB(5),"MINIMIZATION OPTION",Min
1174 PRINT #K0,TAB(5),"OUTPUT OPTION",Iopt3
1175 PRINT #K0,TAB(5),"TIE-BREAKER OPTION",Iopt2
1176 PRINT #K0,TAB(5),"STAGE 1 INPUT OPTION",Iopt1

```

```

1178 REM
1180 Incrmt=1
1190 READ Nstage,Lowd,Maxd,Lows,Maxs,Inc
1200 IF Inc>0 THEN Incrmt=Inc
1202 PRINT #Ko,LIN(1),TAB(5),"NUMBER OF STAGES",Nstage
1203 PRINT #Ko,TAB(5),"RANGE OF DECISION IS",Lowd," TO ",Maxd
1204 PRINT #Ko,TAB(5),"RANGE OF STATES IS",Lows," TO ",Maxs
1205 PRINT #Ko,TAB(5),"INCREMENT",Incrmt
1210 READ Ncolsp,Ncolsm,Npd,Nps,Nrowsc,Ncolsc
1211 IF Iopt3<0 THEN 1220
1212 PRINT #Ko,LIN(1),TAB(5),"NO OF COLS IN 'P' MATRIX",Ncolsp
1213 PRINT #Ko,TAB(5),"NO OF COLS IN 'M' MATRIX",Ncolsm
1214 PRINT #Ko,TAB(5),"NO OF ELEM IN 'MPD'",Npd
1215 PRINT #Ko,TAB(5),"NO OF ELEM IN 'MPS'",Nps
1216 PRINT #Ko,TAB(5),"NO OF ROWS IN 'C' MATRIX",Nrowsc
1217 PRINT #Ko,TAB(5),"NO OF COLS IN 'C' MATRIX",Ncolsc
1219 REM
1220 IF Ncolsp<=0 THEN 1280
1225 FOR I=1 TO Nstage
1230 FOR J=1 TO Ncolsp
1235 READ P(I,J)
1240 NEXT J
1245 NEXT I
1247 IF Iopt3<0 THEN 1280
1250 PRINT #Ko,LIN(1)
1251 FOR I=1 TO Nstage
1252 IF Ncolsp>=4 THEN 1260
1253 FOR J=1 TO Ncolsp
1254 PRINT #Ko,TAB(5),"P(",I,",",J,")=",P(I,J),
1255 NEXT J
1256 GOTO 1274
1260 FOR J=1 TO Ncolsp STEP 4
1261 PRINT #Ko,TAB(5),"P(",I,",",J+0,")=",P(I,J+0),
1262 PRINT #Ko,TAB(5),"P(",I,",",J+1,")=",P(I,J+1),
1263 PRINT #Ko,TAB(5),"P(",I,",",J+2,")=",P(I,J+2),
1264 PRINT #Ko,TAB(5),"P(",I,",",J+3,")=",P(I,J+3)
1270 NEXT J
1274 PRINT #Ko,LIN(1)
1275 NEXT I
1280 IF Nrowsc<=0 THEN 1350
1282 IF Ncolsc<=0 THEN 1350
1285 FOR I=1 TO Nrowsc
1286 FOR J=1 TO Ncolsc
1290 READ C(I,J)
1295 NEXT J
1296 NEXT I
1300 IF Iopt3<0 THEN 1350
1305 PRINT #Ko,LIN(1)
1310 FOR I=1 TO Nrowsc
1312 IF Ncolsm>=4 THEN 1322
1314 FOR J=1 TO Ncolsc
1316 PRINT #Ko,TAB(5),"C(",I,",",J,")=",C(I,J),
1318 NEXT J
1320 GOTO 1335

```

```

1322 FOR J=1 TO Ncolsc STEP 4
1325 PRINT #Ko,TAB(5),"C(",I,"",J+0,"")=",C(1,J+0),
1326 PRINT #Ko,TAB(5),"C(",I,"",J+1,"")=",C(1,J+1),
1327 PRINT #Ko,TAB(5),"C(",I,"",J+2,"")=",C(1,J+2),
1328 PRINT #Ko,TAB(5),"C(",I,"",J+3,"")=",C(1,J+3)
1330 NEXT J
1335 PRINT #Ko,LIN(1)
1340 NEXT I
1345 REM
1350 IF Ncolsm<=0 THEN 1410
1355 FOR I=1 TO Nstage
1357 FOR J=1 TO Ncolsm
1360 READ M(1,J)
1365 NEXT J
1367 NEXT I
1368 IF Iopt3<0 THEN 1410
1369 PRINT #Ko,LIN(1)
1370 FOR I=1 TO Nstage
1371 IF Ncolsm>=4 THEN 1380
1372 FOR J=1 TO Ncolsm
1375 PRINT #Ko,TAB(5),"M(",I,"",J,"")=",M(I,J),
1377 NEXT J
1378 GOTO 1395
1380 FOR J=1 TO Ncolsm STEP 4
1382 PRINT #Ko,TAB(5),"M(",I,"",J+0,"")=",M(I,J+0),
1383 PRINT #Ko,TAB(5),"M(",I,"",J+1,"")=",M(I,J+1),
1384 PRINT #Ko,TAB(5),"M(",I,"",J+2,"")=",M(I,J+2),
1385 PRINT #Ko,TAB(5),"M(",I,"",J+3,"")=",M(I,J+3)
1390 NEXT J
1395 PRINT #Ko,LIN(1)
1400 NEXT I
1405 REM
1410 IF Npd<=0 THEN 1610
1415 FOR I=1 TO Npd
1417 READ Mpd(I)
1420 NEXT I
1430 IF Iopt3<0 THEN 1610
1432 PRINT #Ko,LIN(1)
1435 FOR I=1 TO Npd
1437 PRINT #Ko,TAB(5),"MPD(",I,"")=",Mpd(I),
1440 NEXT I
1450 N=1
1460 Kmm=1
1470 FOR I=1 TO Npd
1480 IF Mpd(Kmm)<0 THEN 1520
1490 Lpd(N)=Kmm
1500 Kmm=Kmm+1
1510 GOTO 1540
1520 Lpd(N)=-(Kmm+1)
1530 Kmm=Kmm-Mpd(Kmm)
1540 IF Kmm<=Npd THEN 1560
1550 PRINT #Ko,"$$$$$ ERROR CONDITION 1  $$$$$$"
1555 GOTO 60000
1560 N=N+1

```

REM

REM

```

1570 IF N>Nstage THEN 1610
1580 Kmm=Kmm+1
1590 NEXT I
1600 REM
1610 IF Np<=0 THEN 1800
1620 FOR I=1 TO Nps
1622 READ Mps(I)
1624 NEXT I
1628 IF Iopt3<0 THEN 1800
1629 PRINT #Ko,LIN(1)
1630 FOR I=1 TO Nps
1635 PRINT #Ko,TAB(5),"MPS(",I,")=",Mps(I),
1637 NEXT I
1640 N=1
1650 Kmm=1
1660 FOR I=1 TO Nps
1670 IF Mps(Kmm)<0 THEN 1710
1680 Lps(N)=Kmm
1690 Kmm=Kmm+1
1700 GOTO 1730
1710 Lps(N)=- (Kmm+1)
1720 Kmm=Kmm-Mps(Kmm)
1730 IF Kmm<=Nps THEN 1750
1740 PRINT #Ko,"$$$$$ ERROR CONDITION 2  $$$ $"
1745 GOTO 60000
1750 N=N+1
1760 IF N>Nstage THEN 1800
1770 Kmm=Kmm+1
1780 NEXT I
1790 REM
1800 IF Iopt2<=1 THEN 1850
1802 FOR I=1 TO Nstage
1804 READ Mtb(I)
1806 NEXT I
1840 GOTO 1880
1850 FOR I=1 TO Nstage
1860 Mtb(I)=Iopt2
1870 NEXT I
1880 IF Iopt3<0 THEN 1890
1882 PRINT #Ko,LIN(1)
1884 FOR I=1 TO Nstage
1886 PRINT #Ko,TAB(5),"MTB(",I,")=",Mtb(I)
1888 NEXT I
1890 Nns=(Maxs-Lows+Incrmt)/Incrmt
1900 IF Nns<=Maxsze THEN 1920
1910 PRINT #Ko,"$$$$$ ERROR CONDITION 3  $$$ $"
1915 GOTO 60000
1920 IF Iopt1<=0 THEN 1940
1922 FOR I=1 TO Nns
1924 READ Fsopt(I)
1926 NEXT I
1928 FOR I=1 TO Nns
1930 READ Idbes1(I)
1932 NEXT I

```

```

1940 COMMAND "M F H HP-[B#1"
1942 PRINT #Ko,LIN(2)
1950 READ Comment$
1952 IF Comment$="END" THEN 2000
1954 PRINT #Ko,TAB(5),Comment$
1956 GOTO 1950
1960 FOR I=1 TO Nns
1970 READ Idbest(I)
1980 NEXT I
2000 REM                                END OF INPUT SECTION                                REM
2010 REM
2020 REM                                OPEN TAPE FILES                                REM
2031 Ntp1=1
2032 ASSIGN "TAPE1" TO #Ntp1
2033 REM                                COMMAND "F F 12 TAPE1"
2034 COMMAND "RL TAPE1"
2035 Ntp2=2
2036 ASSIGN "TAPE2" TO #Ntp2
2037 COMMAND "RE TAPE2"
2039 REM                                REM
2040 Mactp=-1
2050 Mflag=0
2055 Kfile=0
2056 REM                                REM
2060 PRINT #Ko,LIN(3)
2065 PRINT #Ko,TAB(5),"STAGE","STATE","OPT DECISION","OPT VALUEN"
2066 PRINT #Ko,TAB(5),"-----","-----","-----","-----"
2067 PRINT #Ntp2,TAB(5),"STAGE","STATE","DECISION","VALUE","TIED"
2070 PRINT #Ntp2,TAB(5),"-----","-----","-----","-----","-----"
2075 REM                                REM
2080 N1=1
2090 IF lopt1<=0 THEN 2240
2095 Ns=Lows-Incrmt
2100 FOR K=1 TO Nns
2110 Ns=Ns+K*Incrmt
2120 PRINT #Ko,TAB(5),N1,Ns,Idbest(K),Fsopt(K)
2130 NEXT K
2171 FOR L=1 TO Nns
2180 PRINT #Ntp1,Idbest(L),
2190 NEXT L
2220 Comman"M F H TAPE"
2230 N1=2
2235 REM
2240 REM                                START MAIN LOOP                                REM
2245 REM
2250 FOR N=N1 TO Nstage
2260 REM                                INITIALIZE FSBEST(I) & IDBEST(I)                                REM
2270 FOR K=1 TO Nns
2280 Fsbest(K)=Bigm
2290 Idbest(K)=8888
2300 NEXT K
2310 REM                                INITILIZE THE STATE VARIABLE 'NS'                                REM
2320 IF Nps>0 THEN 2410
2330 Ns=Lows
2340 Hs=Maxs

```

```

2350 IF N<>Nstage THEN 2570
2380 IF Nps=-1 THEN Ns=Maxs
2390 IF Nps=-2 THEN Hs=Lows
2400 GOTO 2570
2410 Locs=Lps(N)
2420 IF Locs<0 THEN 2500
2430 Ns=Mps(Locs)
2440 IF Ns>Maxs THEN 2455
2450 IF Ns<Lows THEN 2460
2455 PRINT #Ko,"$$$$$ ERROR CONDITION 5  $$$ $"
2458 GOTO 60000
2460 Hs=Mps(Locs+1)
2470 IF Hs>Maxs THEN 2490
2480 IF Hs<Lows THEN 2470
2490 PRINT #Ko,"$$$$$ ERROR CONDITION 6  $$$ $"
2496 GOTO 60000
2500 Locs=-Locs
2510 Ns=Mps(Locs)
2520 IF Ns>Maxs THEN 2535
2530 IF Ns<Lows THEN 2540
2535 PRINT #Ko,"$$$$$ ERROR CONDITIION 7  $$$ $"
2536 GOTO 60000
2540 K=Locs-1
2550 Lshigh K-Mps(K)
2570 REM
2571 I=(Ns-Lows+Incrmt)/Incrmt
2580 ldbest(I)=8888
2590 IF Npd>0 THEN 2630
2600 Nd=Lowd
2610 Hd=Maxd
2620 GOTO 2790
2630 Locd=Lpd(N)
2640 IF Locd<0 THEN 2720
2650 Nd=Mpd(Locd)
2660 IF Nd>Maxd THEN 2675
2670 IF Nd<Lowd THEN 2680
2675 PRINT #Ko,"$$$$$ ERROR CONDITION 9  $$$ $"
2676 GOTO 60000
2680 Hd=Mpd(Locd+1)
2690 IF Hd>Maxd THEN 2710
2700 IF Hd<Nd THEN 2790
2710 PRINT #Ko,"$$$$$ ERROR CONDITION 10  $$$ $"
2715 GOTO 60000
2720 Locd=-Locd
2730 Nd=Mpd(Locd)
2740 IF Nd>Maxd THEN 2754
2750 IF Nd<Lowd THEN 2760
2754 PRINT #Ko,"$$$$$ ERROR CONDITION 11  $$$ $"
2755 GOTO 60000
2760 K=Locd-1
2770 Ldhigh K-Mpd(K)
2790 REM
2791 IF Funct<=0 THEN 2800
2792 IF Funct<=9 THEN 2805

```

1 150

1 175

1 200

1 250

1 275

1 300

1 300


```

2800 PRINT #Ko, "##### ERROR CONDITION 14 #####"
2802 GOTO 60000
2805 Kn=N                                     | 400
2806 Ks=Ns
2807 Kx=Nd
2810 ON Funct GOSUB 5520,5660,5760,5800,5930,6310,6430,6570,6760
2815 REM
2820 Itied=0                                 | 310
2830 IF Fnsxn=Fsbest(I) THEN Itied=-1
2835 PRINT #Ntp2,TAB(5),N,Ns,Nd,Fnsxn
2840 IF Min>0 THEN 3080
2850 REM                                     MAXIMIZING LOGIC REM
2860 IF Mtb(N)<=0 THEN 2890
2870 IF Fnsxn>=Fsbest(I) THEN 2920
2880 GOTO 2950
2890 IF Fnsxn<=Fsbest(I) THEN 2950         | 325
2900 REM                                     UPDATE VALUE AND DECISION REM
2910 Fsbest(I)=Fnsxn                       | 340
2930 ldbest(I)=Nd
2940 REM                                     INCREMENT THE DECISION VARIABLE 'ND' REM
2950 IF Itied=0 THEN 2980                 | 350
2960 Ix=ldbest(I)
2970 ldbest(I)=-ABS(Ix)
2980 IF Npd<=0 THEN 3050                 | 355
2990 IF Lpd(N)>=1 THEN 3050
3000 Locd=Locd+1
3010 IF Locd>Ldhigh THEN 3180
3020 Nd=Mpd(Locd)
3030 IF Nd<=Maxd THEN 2790
3034 IF Nd>=Lowd THEN 2790
3040 PRINT #Ko, "##### ERROR CONDITION 12 #####"
3045 GOTO 60000
3050 Nd=Nd+Incrmt                         | 375
3060 IF Nd<=Hd THEN 2790
3070 GOTO 3180
3080 REM                                     MINIMIZING LOGIC REM
3090 IF Mtb(N)<=0 THEN 3120               | 380
3100 IF Fnsxn<=Fsbest(I) THEN 2920
3110 GOTO 2950
3120 IF Fnsxn<Fsbest(I) THEN 2920         | 390
3130 GOTO 2950
3180 IF lopt3>0 THEN 3220                 | 500
3190 IF N<>Nstage THEN 3310
3220 REM                                     OUTPUT OPTIONAL EXCEPT FOR LAST STAGE REM
3221 Ix=ldbest(I)
3230 Itied$=" "
3240 IF Ix>=0 THEN 3280
3250 Ix=-Ix
3260 Itied$="*"
3270 Mflag=1
3280 PRINT #Ko,TAB(5),N,Ns,Ix,Fsbest(I),Itied$ | 505
3290 IF N=Nstage THEN 3580
3310 REM                                     INCREMENT THE STATE VARIABLE 'NS' REM
3320 PRINT #Ntp2,LIN(1)

```

```

3321 IF Nps<=0 THEN 3390
3322 IF Lps(N)>=1 THEN 3390
3330 Locs=Locs+1
3340 IF Locs>Lshigh THEN 3410
3350 Ns=Mps(Locs)
3360 IF Ns>Maxs THEN 3380
3370 IF Ns<Lowest THEN 3380
3375 GOTO 2570
3380 PRINT #Ko,"$$$$$ ERROR CONDITION 8  $$$$$"
3382 GOTO 60000
3390 Ns=Ns+Incrmt
3400 IF Ns<=Hs THEN 2570
3410 IF lopt3>0 THEN PRINT #6,LIN(1)
3420 IF N=Nstage THEN 4430
3430 REM                                SAVE CURRENT STAGE DECISION ON TAPE1    REM
3471 FOR L=1 TO Nns
3480 PRINT #Ntp1,lbest(L),
3490 NEXT L
3515 COMMAND "M F H TAPE1"
3530 REM                                UPDATE FSOPT(I) FOR USE AT NEXT STAGE    REM
3531 FOR K=1 TO Nns
3540 Fsopt(K)=Fsbest(K)
3550 NEXT K
3560 GOTO 4430
3580 REM                                RECOVER AND OUTPUT OF OPTIMAL DECISIONS    REM
3581 N2=Nstage
3590 Js=Ns
3600 M2=1
3610 Kdec(M2)=Ix
3620 Nty=0
3630 IF lbest(I)>=0 THEN 3660
3640 Nty=Nty+1
3650 Kties(Nty)=N2
3655 REM
3660 IF N2=1 THEN 4230
3670 Jx=Kdec(M2)
3680 IF Jx=8888 THEN 3310
3690 IF Funct<=0 THEN 3310
3700 IF Funct>9 THEN 3310
3710 Kn=N2
3720 Ks=Js
3730 Kx=Jx
3740 REM                                LOOKUP THE RECOVERY FUNCTION    REM
3750 ON Funct GOSUB 5240,5210,5210,3310,3310,5240,5360,5330,5270
3760 Js=News
3770 IF Js=8888 THEN 3310
3780 IF N2<>Nstage THEN 3870
3790 IF Moetp>0 THEN 3910
3830 REM                                PUT TAPE IN READING MODE    REM
3840 COMMAND "F F -1 TAPE1"
3845 Moetp=-Moetp
3850 GOTO 3940
3860 REM                                BACKSPACE TAPE TO POSITION IN FRONT OF NEXT STAGE    REM
3870 COMMAND "F F -1 TAPE1"

```

```

3900 REM REM
3910 COMMAND "F F -1 TAPE1" I 620
3940 REM READ OPT DECISION OF PREVIOUS STAGE REM
3950 FOR L=1 TO Nns I 623
3960 READ #Ntp1,Idbest(L)
3970 NEXT L
3980 Kss=(Js-Lows+Incrmt)/Incrmt
3990 N2=N2-1
4000 M2=M2+1
4020 Kdec(M2)=Idbest(Kss)
4030 IF Kdec(M2)>=0 THEN 3660
4040 Kdec(M2)=-Kdec(M2)
4070 GOTO 3640
4220 REM REM
4230 N3=Nstage+1 I 850
4231 PRINT #Ko,LIN(2)
4235 PRINT #Ko,TAB(15),"STAGE","DECISION"
4236 PRINT #Ko,TAB(15),"-----","-----"
4240 FOR K=1 TO Nstage
4250 PRINT #Ko,TAB(15),N3-K,Kdec(K)
4260 NEXT K
4270 PRINT #Ko,LIN(2)
4280 IF Nty=0 THEN 4340
4285 PRINT #Ko,TAB(5),"ALTERNATE OPTIMAL DECISIONS EXIST AT STAGE'S "
4290 FOR K=1 TO Nty
4300 PRINT #Ko,TAB(5),Kties(K)
4310 NEXT K
4320 PRINT #Ko,LIN(1)
4340 REM REPOSITION TAPE REM
4395 COMMAND "F E D L"
4400 COMMAND "F F -1 TAPE1"
4410 GOTO 3310
4420 REM
4430 PRINT #Ntp2,LIN(1)
4435 NEXT N
4440 REM END OF THE MAIN LOOP FOR EACH STAGE REM
4450 IF Mflag<=0 THEN 60000
4460 PRINT #6,TAB(5)," ** AFTER THE OPTIMUM DECISION INDICATES THAT"
4470 PRINT #6,TAB(5)," ALTERNATE OPTIMAL DECISIONS EXIST "
4480 PRINT #6,TAB(5)," THE TIE-BREAKER OPTION CONTROLS SELECTION"
4490 PRINT #6,TAB(5)," OF THE LOWEST OR HIGHEST DECISION "
4540 GOTO 60000
4550 REM 1 2
4560 REM END OF PROGRAM REM
4570 REM LIST 4340
4600 REM FUNCTION LOCATE (IJ) REM
4610 Locat=(Ij-Lows+Incrmt)/Incrmt
4620 RETURN
4700 REM FUNCTION LOCATD (MN,MD) REM
4702 IF Npd>0 THEN 4708
4704 L1=Lowd
4706 GOTO 4714
4708 IF Lpd(Mn)<=0 THEN 4718

```

```

4710 Kk=Lpd(Mn)
4712 LJ=Mpd(Kk)
4714 Locatd=(Md-LJ+Incrmt)/Incrmt
4716 RETURN
4718 Kk=-Lpd(Mn)
4720 Locatd= 0
4722 Jdec=Mpd(Kk)
4724 Locatd=Locatd+1
4726 IF Jdec=Md THEN 4716
4728 Kk=Kk+1
4730 IF Kk<=Ldhigh THEN 4722
4732 PRINT #Ko,"$$$$$ ERROR CONDITION 13  $$$%"
4734 GOTO 60000
4800 REM                                FUNCTION FSTAR (NEW,VL,VH,V1)                                REM
4802 IF New<Lows THEN 4814
4804 IF New>Maxs THEN 4818
4806 IF N=1 THEN 4822
4808 Kk=(New-Lows+Incrmt)/Incrmt
4810 Fstar=Fsopt(Kk)
4812 RETURN
4814 Fstar=V1
4816 RETURN
4818 Fstar=Vh
4820 RETURN
4822 Fstar=V1
4824 RETURN
38600 Jx=Kdec(M2)
60000 REM                                REM
60010 COMMAND "RE TAPE 1"
60011 COMMAND "RE TAPE 2"
60020 COMMAND "M F H HP-IB 1"
60030 PRINT #Ko,LIN(3)
60040 PRINT #Ko,"", "*****"
60060 PRINT #Ko,"", "**"
60070 PRINT #Ko,"", "**                                END OF TASK                                "**
60080 PRINT #Ko,"", "**"
60090 PRINT #Ko,"", "*****"
60100 END

```

```

5270 REM                                FUNCTION NEWS3 (KN,KS,KX)                                REM
5280 IF Ncolsm>1 THEN 5305
5290 News=Ks-(Kx)*M(Kn,1)
5300 RETURN
5305 Mn=Kn
5306 Md=Kx
5308 GOSUB 4700
5310 News=Ks-M(Kn,Locatd)
5320 RETURN
6760 REM                                FUNCTION RECUR9 (KN,KS,KX)                                REM
6770 REM                                GENERAL SOLUTION FOR SYSTEM RELIABILITY                                REM
6800 IF Ncolsp=1 THEN 6840
6802 Mn=Kn
6804 Md=Kx
6810 COSUB 4700
6820 Fnsxn=P(Kn,Locatd)
6830 GOTO 6850
6840 Fnsxn=1-(1-P(Kn,1))^(Kx)
6850 V1=Bigm
6860 Vh=Bigm
6870 V1 1
6875 GOSUB 5280
6877 New=News
6880 GOSUB 4800
6890 Fnsxn=Fnsxn*Fstar
6900 RETURN

```

Appendix B

Example 1 Inputs and Result

```

7000 RLM                                INPUT DATA                                REM
7100 RLM ENTER DATA FUNCT, MIN, IOPT3, IOPT2, IOPT1
7110 DATA 9, 0, 0, 0, 0
7150 REM ENTER DATA NSTAGE, LOWD, MAXD, LOWEST, MAXS, INC
7151 DATA 4, 1, 3, 0, 45, 1
7200 REM ENTER DATA NCOLSP, NCOLSM, NPD, NPS, NROWSC, NCOLSC
7210 DATA 3, 3, 0, -1, 0, 0
7250 REM ENTER DATA ( (P(I,J), J=1,NCOLSP), I=1,NSTAGE )
7251 DATA 60, 75, 85
7252 DATA 40, 65, 80
7253 DATA 70, 90, 95
7254 DATA 50, 60, 80
7350 REM ENTER DATA ( (M(I,J), J=1,NCOLSM), I=1,NSTAGE )
7351 DATA 6, 11, 15
7352 DATA 10, 16, 22
7353 DATA 5, 10, 14
7354 DATA 8, 13, 17
7400 REM ENTER COMMENTS
7410 DATA "TYPE 9 EXAMPLE 1"
7411 DATA " "
7412 DATA " "
7420 DATA "STAGE.      COMPONENTS OF THE SYSTEM"
7421 DATA "STAGE      UNALLOCATED RESOURCES (DOLLARS)"
7422 DATA "DECISION    NUMBER OF PARALLEL UNITS OF THE COMPONENT INSTALLED"
7423 DATA "VALUE      .  PROBABILITY OF THE SYSTEM WILL FUNCTION SUCCESS"
7424 DATA "END"

```

TYPE 9 EXAMPLE 1

STAGE COMPONENTS OF THE SYSTEM
 STATE UNALLOCATED RESOURCES (DOLLARS)
 DECISION NUMBER OF PARALLEL UNITS OF THE COMPONENT INSTALLED
 VALUE PROBABILITY OF THE SYSTEM WILL FUNCTION (SUCCESS)

STAGE	STATE	OPT DECISION	OPT VALUEN
-----	-----	-----	-----
4	45	1	219375

STAGE	DECISION
-----	-----
4	1
3	2
2	2
1	2

Appendix C

Example 2 Inputs and Result

```

7000 RLM ENTER DATA FUNC1, MIN, IOPT3, IOPT2, IOPT1
7001 DATA 9, 0, 0, 0, 0
7002 RLM ENTER DATA NSTAGE, LOWD, MAXD, LOWEST, MAXS, INC
7003 DATA 4, 1, 4, 0, 45, 1
7004 RLM ENTER DATA NCOLSP, NCOLSM, NPD, NPS, NROWSC, NCOLSC
7005 DATA 1, 4, 0, -1, 0, 0
7006 RLM ENTER DATA ( (P(I,J), J=1,NCOLSP), I=1,NSTAGE )
7007 DATA 60
7008 DATA 40
7009 DATA 70
7010 DATA 50
7011 RLM ENTER DATA ( (M(I,J), J=1,NCOLSM), I=1,NSTAGE)
7012 DATA 6, 11, 15, 18
7013 DATA 10, 16, 22, 27
7014 DATA 5, 10, 14, 17
7015 DATA 8, 13, 17, 20
7016 RLM ENTER DATA (MPD(I), I=1,NPD)
7017 RLM ENTER DATA (MPS(I), I=1,NPS)
7019 RLM ENTER DATA (MIB(I), I=1,NSTAGE)
7020 DATA "TYPE 9 EXAMPLE 2"
7021 DATA " "
7022 DATA " "
7023 DATA "STAGE      COMPONENTS OF THE SYSTEM"
7024 DATA "STATE      UNALLOCATED RESOURCES (DOLLARS)"
7025 DATA "DECISION    NUMBER OF PARALLEL UNITS OF THE COMPONENT INSTALLED"
7026 DATA "VALUE       PROBABILITY OF THE SYSTEM WILL FUNCTION SUCCESS"
7027 DATA "END"

```


TYPE 9 EXAMPLE 2

STAGE COMPONENTS OF THE SYSTEM
 STATE UNALLOCATED RESOURCES (DOLLARS)
 DECISION NUMBER OF PARALLEL UNITS OF THE COMPONENT INSTALLED
 VALUE PROBABILITY OF THE SYSTEM WILL FUNCTION SUCCESS

STAGE	STATE	OPT DECISION	OPT VALUEN
-----	-----	-----	-----
4	45	2	28224

STAGE	DECISION
-----	-----
4	2
3	1
2	2
1	2

A Computerized Life-Cycle Cost Methodology for Engineering Analysis

R D Hughes
DSN Engineering Section

Life-Cycle Costing (LCC) is an essential selection criterion in making economical engineering decisions about alternative routes in design or investments. A discussion of Life-Cycle Costing (LCC) concepts is presented, along with a selected calculation procedure. A computer program (LCOMP) was written in FORTRAN to perform that calculation procedure. The program details are discussed, a sample calculation is presented, and a listing of the program is included.

I. Introduction

Engineers have been using various economic criteria in making decisions between alternate designs or investments. Life-Cycle Costing (LCC) is one such method of economic evaluation which takes into account all relevant costs of any system, or subsystem, over a specified period of time. The LCC procedure makes adjustments for differences in the timing of these costs, taking into account future fuel and non-fuel cost escalation rates and discount effects which reflect the "time value of money."

Since LCC determines the effective cost of a system over a given lifetime, it is used by engineers to choose between alternate facility modifications and upgrades. Subsequently, the economic feasibility of proposed facility modifications can be determined and acceptable modifications can be prioritized. Although this procedure is especially useful when performing cost-saving energy consumption reduction modifications, it can be applied to other decision-making processes involving alternate configurations or designs not related to energy consumption reduction.

Several LCC methodologies appear in the literature (Refs 1-5). These methodologies were reviewed and evaluated with the following criteria in mind:

- (1) The methodology should fit within the guidelines of TDA Standard Practice for LCC, as described in Refs 6 and 7.
- (2) It should be readily adaptable for use on a computer.
- (3) It should be capable of analyzing energy-related and non-energy related projects, new construction projects, and retrofit projects.
- (4) The resulting computer program should be easy to use and operate efficiently on a digital computer.

The intended result of satisfying these criteria was to develop an inexpensive, general engineering tool for system economic evaluation.

The LCC methodology chosen for computerization is the DOE-NBS methodology for the Federal Energy Management

Program (FEMP), as described in Ref 1 This methodology deals with LCC procedures applied to federal facilities energy management and evaluation, taking varying energy costs into account

II. LCC Concepts and Methodology

The selected DOE LCC methodology performs accounting in constant dollars (money always referred to in terms of its value in a chosen baseyear) and discounts future amounts to present value baseyear amounts In this fashion, the rate of inflation is eliminated from the computations since energy price escalation rates relative to inflation are included in the discounting procedure Thus, LCC incorporates initial investment costs, future replacement costs, operation and maintenance costs, and salvage and resale values, adjusting them to a consistent time basis and combining them into a single cost-effectiveness measure that facilitates comparison of alternate projects

The changing value of money over time is controlled by two effects, inflation and "opportunity costs" Future prices which change at the same rate as general price inflation remain unchanged in terms of "constant dollars" Future prices which increase at a rate different than that of inflation must be expressed in terms which reflect that difference Future energy prices are calculated using prescribed escalation rates Tables of energy escalation rates as projected by DOE are provided in Tables 1 through 3 The "opportunity cost of money" reflects the fact that money in hand can be invested to yield a return above the rate of inflation The "discount rate" is a rate of interest corresponding to the opportunity cost. On June 30, 1980, the Energy Security Act was enacted to establish a discount rate of 7% per year. This rate applies to projects which fall under auspices of the FEMP, but is also an acceptable rate for general use

The common time basis used for this methodology is the present, whereby all cash amounts are converted to an equivalent present value If the amount is an annually recurring amount (A) which remains the same in terms of constant dollars, its present value (P) for a period of N years may be expressed by the uniform present worth (UPW) formula

$$P = \frac{A}{d} \left(1 - \frac{1}{(1+d)^N} \right) \quad (1)$$

where d is the discount rate Operation and maintenance (O&M) costs fall into this category The present value of non-annually recurring costs (F) are calculated by the single present worth (SPW) formula

$$P = F \left(\frac{1}{(1+d)^N} \right) \quad (2)$$

Replacement costs and salvage values are examples of non-annually recurring costs. The present value of annually recurring costs which escalate or de-escalate in relation to inflation may be expressed by the modified uniform worth (UPW*) formula

$$P = A_0 \left[\sum_{j=1}^{n_1} \left(\frac{1+e_1}{1+d} \right)^j + \left(\frac{1+e_1}{1+d} \right)^{n_1} \sum_{j=1}^{n_2} \left(\frac{1+e_2}{1+d} \right)^j + \dots + \left(\frac{1+e_1}{1+d} \right)^{n_1} \left(\frac{1+e_2}{1+d} \right)^{n_2} \left(\frac{1+e_{k-1}}{1+d} \right)^{n_{k-1}} \sum_{j=1}^{n_k} \left(\frac{1+e_k}{1+d} \right)^j \right] \quad (3)$$

where

n_k = length of escalation period k

e_k = escalation rate during period k

A_0 = annually recurring escalating amount, evaluated at the beginning of the study period

Equation (3) is usually applied to energy costs The computer program which embodies this methodology (LCOMP) has the escalation rates contained internally

The formula for life-cycle cost becomes

$$\text{Total Life-Cycle Cost (TLCC)} = \text{Investment Cost} - \text{Salvage Value} + \text{Replacement Costs} + \text{Energy Costs} \quad (4)$$

where all costs are expressed in present values TLCC is used for comparison of new designs or comparison of alternatives for a certain system

Other results of the LCC analysis are the following

- (1) Net Life-Cycle Savings (NS) is a comparative quantity which indicates the difference between the TLCC of two candidate system designs

- (2) Savings-to-Investment Ratio (SIR) is the ratio of the savings in total system operating costs to the investment cost required to install or construct the more efficient system.

$$\text{SIR} = (\text{reduction in energy cost} - \text{increase in O\&M costs}) - (\text{increase in initial investment costs} - \text{increase in salvage value} + \text{increase in replacement costs})$$

(5)

SIR is meaningful, for instance, when comparing retrofit projects to existing facilities. Several alternative retrofit systems can be assigned priorities based on SIR. In general, an SIR value greater than 1.0 indicates cost effectiveness and greater values indicate greater cost effectiveness.

- (3) Payback Period (PB) is the amount of time it takes for the cumulative savings to equal the initial investment costs. There are two versions of the payback period. The "discounted payback" (DPB) is calculated taking the time value of money into account, and the "simple payback" (SPB) uses costs which do not take this into account. The general payback formula is

$$\sum_{j=1}^N (\text{Reduction in Energy Costs}_j - \text{Differential O\&M Costs}_j - \text{Differential Replacement Costs}_j) = \text{Differential Initial Investment Costs}$$

(6)

where N = the years to payback, such that the above relation is true

For DPB the costs are yearly amounts in constant dollars converted to present values, for SPB the costs are baseyear amounts and are not discounted. Ordinarily, a shorter PB is desirable. However, PB is not always a useful measure of cost effectiveness since the project with a longer PB can have a greater NS and SIR and actually be more cost effective. PB is generally considered to be a less accurate measure of relative cost effectiveness than the other LCC results, but it is sometimes a necessary indicator that a system fits within certain requirements.

III. LCC Applications

The four main selection criteria or corollaries of the LCC analysis, namely TLCC, NS, SIR, and PB, will have varying levels of importance based upon the application of the LCC analysis. In general, a project is comparatively cost effective if

- (1) The TLCC of the proposed project is less than that of the alternatives
- (2) The NS from the project is greater than zero
- (3) The SIR is greater than 1.0
- (4) The payback period is shorter than the project's expected life

For choosing among alternative designs, the TLCC is generally the best indicator. For designing and sizing projects the choice should be that which minimizes the TLCC and maximizes the NS. For ranking retrofit projects to give priority to the most cost-effective projects, the SIR is the most useful result.

IV. Computer Program (LCOMP) Description

The LCC methodology described above was used in the program LCOMP (Life-Cycle Costing Computer Program). The calculation procedure is similar to that in the NBS-DOE manual, Ref. 1, where an LCC computer program is also presented. LCOMP differs from this program in several ways:

- (1) LCOMP has built-in energy escalation rate tables and energy base-price tables, freeing the user from looking up and inputting these values.
- (2) LCOMP uses a NAMELIST input instead of an interactive input which allows storage, modification, and reuse of input files.
- (3) LCOMP uses the year-by-year method of calculating present value energy costs, which is necessary if the quantity or type of energy is expected to change over the project study period and if cash flow tables are presented as output.
- (4) LCOMP is written in FORTRAN instead of BASIC.

LCOMP calculates and prints out (1) present values (in baseyear dollars) of total cost of each energy type used, (2) total of all energy costs, (3) total annually recurring O&M costs, (4) total replacement costs, (5) total salvage values, and (6) the system total life cycle cost. In addition, if two systems are being compared, the NS, SIR, and PB are printed, along with a year-by-year cash flow summary.

LCOMP was written in FORTRAN V and currently is being run on the UNIVAC 1100/81 computer system. A block diagram of the program logic is presented in Appendix A and a source listing is presented in Appendix B.

The program input consists of two segments. The first segment is the NAMELIST "\$LCCIN," which contains most of

the case-dependent data. The second segment is the Blockdata subprogram "BLKDAT," which contains energy cost escalation data and baseyear energy prices. These inputs are explained in detail in Appendix C.

An important aspect of operating LCOMP is that the energy price and escalation rate tables are based on a particular baseyear's currency (the tables now in Blockdata are in 1980 dollars). This means that the analysis will be done in the baseyear currency, even if the project begins at some future year. If cost inputs cannot be estimated in baseyear dollars, a discrepancy of one or two years may not cause significant errors, especially if a comparison between two systems is the important objective of the LCC analysis. However, if more than a few years difference exists between currency baseyears, an attempt should be made to update the baseyear tables.

The DOE methodology makes use of an additional "social benefit adjustment factor," which is intended to reflect the value to the nation of conserving non-renewable energy sources. The presently recommended procedure is to reduce the investment costs of a new or retrofit system by 10% to take this social factor into account. However, this adjustment is not performed in the LCOMP program. Its use is left to the discretion of the user, and it must be performed on the input data if desired.

V. A Sample Problem

A four-page printed output for a sample problem is presented in Appendix D to illustrate the LCOMP output format. Figure D-1 consists of the project description and non-fuel costs, and Fig. D-2 shows a summary of annual fuel consumption. Figure D-3 is LCOMP printout of a cash flow analysis which shows year-by-year costs discounted to present value 1980 dollars. The line for year 0 reflects the investment cost for each system, where the investment cost for System 2 in this case is the current salvage value. The final column, "System 1 vs. System 2 Cumulative Savings," represents the cumulative cost of System 2 minus the cumulative cost of System 1. Thus, in this example, the cumulative cost of the existing system exceeds the cumulative cost of the retrofit system in year 7, and the savings then becomes positive. This quantity represents a discounted payback period with energy price escalation included.

Figure D-4 shows analysis results. Total costs for the study period are itemized and their sum represents the system life-

cycle cost. The net savings is the difference between the life-cycle costs. Note that net savings agrees with the final value in the cumulative savings column in Fig. D-3. The savings-to-investment ratio is in terms of present values. The simple payback analysis is based on a non-discounted cumulative savings compared to increased investment cost. The discounted payback is usually preferred over the simple payback method, but this choice depends on the mode of analysis.

VI. Use of LCOMP Within TDA Guidelines

The TDA methodology deals with issues concerning when to perform an LCC analysis, what applicability the analysis has, when it is actually required, LCC team roles, how to perform the different classes of cost estimates, and what type of cost adjustment procedure is to be used. It is only the last concern which affects LCC calculation details. The basic TDA formulations of the LCC equations for TLCC, NS, SIR, DPB are analogous to those of the DOE procedure.

The cost adjustment procedure is prescribed according to the application of the LCC analysis as abstracted from Ref. 7.

LCC analysis application	Procedure
Design selection	No adjustment. Sponsor may require sensitivity analysis using net-discounting.
Capability planning	No adjustment.
Functional requirement negotiation	No adjustment.
Budget planning	Inflate. Sponsor may require sensitivity analysis using no adjustment.

The "no adjustment" procedure is an LCC analysis done in constant dollars with no inflation or discounting adjustments included. The "net discounting" procedure attempts to take inflation and discounting into account in one parameter by assuming that the discount rate tends to exceed the inflation rate by 2% (Ref. 5). Both of these cost adjustment procedures may be handled directly by the program LCOMP. The no adjustment procedure would be computed using a discount rate of zero and the net discounting procedure would use a discount rate of 2%.

The inflation procedure, used in making budget projections, gives the expected expenditure in dollars that will be required.

in the year in which the activity will occur. Again, LCOMP can be used to perform an LCC analysis with zero discount, but the cash flow quantities will have to be inflated in a separate calculation.

The TDA procedure calls for the application of a management contingency factor and a composite burden factor which are chosen by the analysis coordinator. These factors may be applied to costs before inputting data to LCOMP.

VII. Summary

LCOMP is an easy-to-use computer program written by the Advanced Technology and Conservation Engineering Group, DSN Engineering Section, which calculates Life-Cycle Cost on a discounted present value approach. Although it follows DOE-defined methodology, LCOMP is applicable to any LCC problem for alternate systems comparison whether or not it is dealing with energy conversion/conservation.

References

1. Ruegg, R. T., "Life-Cycle Costing Manual for the Federal Energy Management Program," *NBS Handbook 135*, Dec 1980.
2. Brown, R. J., and Yanuck, R., *Life-Cycle Costing: A Practical Guide for Energy Managers*, Fairmont Press, Inc., Atlanta, Georgia, 1980.
3. Baltimore Aircoil Co., Inc., "A Guide to Life-Cycle Costing," Bulletin S602/1-0, Baltimore, Maryland, 1977.
4. Hollander, G., "Life Cycle Cost: A Concept in Need of Understanding," *Professional Engineer Magazine*, June 1976.
5. Eisenberger, I., Remer, D. S., and Lorden, G., "The Role of Interest and Inflation Rates in Life-Cycle Cost Analysis," in *The Deep Space Network Progress Report 42-43*, pp. 105-109, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1978.
6. McKenzie, M., *DSN Life-Cycle Cost Analysis Handbook*, JPL Document 890-119, Rev. A, Jan. 15, 1981.
7. McKenzie, M., "Life-Cycle Cost Analysis," *TDA Standard Practice*, JPL Document 810-23, Sept. 15, 1981.

Table 1. Energy price escalation rates mid-1980 to mid-1985^a (percentage change compounded annually)

Fuel type	DOE region										U S average
	1	2	3	4	5	6	7	8	9	10	
Residential sector											
Electricity	-0 02	-0 02	-0 02	-0 02	-0 01	-0 02	-0 02	-0 02	-0 01	-0.02	-0 02
Natural gas	1 78	1 78	1 77	1.75	1 75	1.76	1 75	1 76	1 77	1.78	1.76
Distillate	3 38	3 39	3 38	3 39	3 37	3 38	3 38	3 39	3 37	3 37	3 38
Commercial sector											
Electricity	-0 01	-0 02	-0 01	-0 01	-0 02	-0 01	-0 01	-0.02	-0 01	-0 02	-0 02
Natural gas	1 77	1 76	1 75	1 73	1 74	1.74	1.73	1 75	1 76	1.76	1.75
Distillate	3 38	3 39	3 38	3.38	3 39	3 38	3 40	3 38	3 38	3 38	3 39
Residual	7 53	7 52	7 52	7 53	7 55	7 52	7 55	7 51	7 50	7.57	7 53
Industrial sector											
Electricity	-0 01	-0 03	-0 01	-0 02	-0.01	-0 02	-0 02	-0 01	-0 01	-0 05	-0 03
Natural gas	1 80	1 76	1 78	1 76	1.77	1 79	1 74	1 80	1.79	1.79	1 75
Distillate	3 39	3 40	3 39	3 39	3.39	3 38	3 40	3.37	3 38	3 38	3 38
Residual	7 54	7 53	7 53	7 53	7 54	7 53	7 54	7.53	7 51	7.55	7 53
Coal	9 62	9 51	9 63	9 58	9 49	9 62	9 50	9 30	9 56	9 56	9 55

^aDerived from DOE 1980 and 1985 price forecasts (Ref. 1)

Table 2 Energy price escalation rates mid-1985 to mid-1990^a (percentage change compounded annually)

Fuel type	DOE region										U S average
	1	2	3	4	5	6	7	8	9	10	
Residential sector											
Electricity	-0 02	-0 61	0 87	1 72	1 04	1 53	-0 59	-2 73	0 47	3 85	0 85
Natural gas	3 33	2 74	3 15	2 38	2 84	4 53	3 43	3 95	1 44	3 86	2 92
Distillate	2 81	2 80	2 74	2 71	2 91	2 83	2 94	2 82	2 97	2 97	2 85
Commercial sector											
Electricity	-0 19	-0 64	0 89	1 67	1 07	1 62	-0 63	-2 96	0 43	3 97	0 73
Natural gas	3 88	3 20	3 60	2 82	3 15	5 26	3 85	4 22	1 66	4 50	3 49
Distillate	2 91	2 88	2 89	2 88	2 99	2 93	3 01	2 91	3 09	3 09	2 94
Residual	2 66	2 68	2 52	2 71	2 70	2 71	2 69	2 67	2 84	2 92	2 61
Industrial sector											
Electricity	-0 23	-0 98	1 19	2 20	1 47	2 03	-0 78	-4 05	0 55	7 89	1 32
Natural gas	3 81	4 46	8 89	8 82	6 98	5 88	11 74	4 95	3 26	6 20	6 64
Distillate	3 47	2 88	2 85	2 85	2 99	2 94	3 02	2 90	3 09	3 09	2 93
Residual	3 53	2 60	2 58	2 73	2 71	2 71	2 70	2 69	2 87	2 86	2 68
Coal	1 47	1 65	1 97	1 69	1 67	1 45	1 76	0 00	1 36	2 39	1 66

^aDerived from DOE mid-term energy price forecasts (Ref. 1)

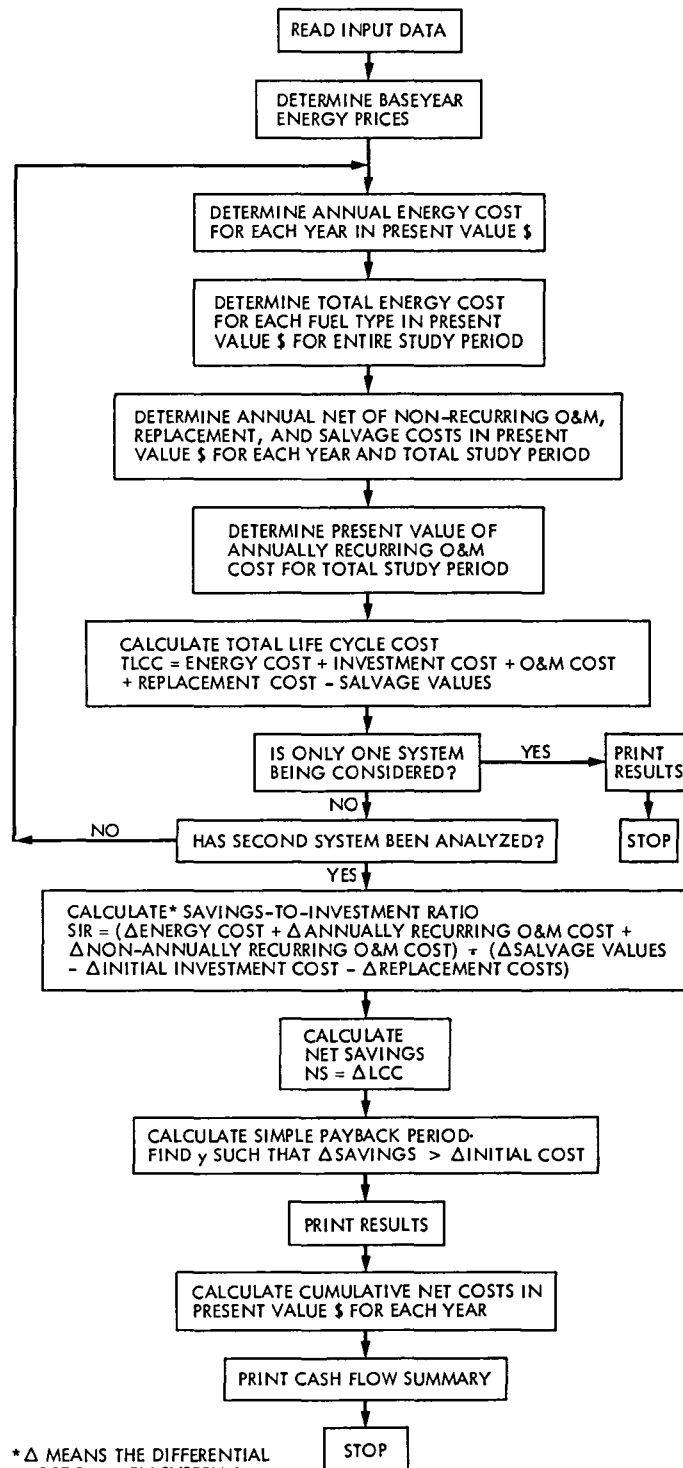
Table 3. Energy price escalation rates mid-1990 to mid-1995 and beyond^a (percentage change compounded annually)

Fuel type	DOE region										U S average
	1	2	3	4	5	6	7	8	9	10	
Residential sector											
Electricity	-3.55	-0.42	-0.36	0.48	-0.13	-0.26	-0.02	-2.47	-2.35	1.10	-0.57
Natural gas	1.92	1.68	1.56	1.01	0.97	2.33	1.23	2.36	0.26	-3.09	1.24
Distillate	3.97	3.95	3.89	3.86	4.10	4.00	4.13	4.06	4.13	4.13	4.01
Commercial sector											
Electricity	-3.60	-0.44	-0.37	0.48	-0.14	-0.28	-0.03	-2.70	-2.21	1.13	-0.59
Natural gas	2.16	1.93	1.80	1.18	1.11	2.67	1.38	2.49	0.29	-3.58	1.39
Distillate	4.08	4.09	4.07	4.06	4.20	4.12	4.22	4.18	4.27	4.27	4.09
Residual	4.47	4.41	4.26	4.50	4.56	4.56	4.54	4.60	4.68	4.82	4.43
Industrial sector											
Electricity	-4.45	-0.68	-0.50	0.63	-0.19	-0.34	-0.03	-3.91	-2.78	1.97	-0.43
Natural gas	4.72	4.60	-0.95	3.41	3.92	2.89	9.60	2.84	0.93	3.99	3.35
Distillate	4.08	4.10	4.02	4.02	4.20	4.13	4.22	4.16	4.27	4.27	4.12
Residual	4.43	4.35	4.28	4.52	4.58	4.57	4.53	4.64	4.68	4.71	4.55
Coal	-3.47	1.00	1.12	0.79	0.98	0.60	0.77	0.62	0.94	-2.87	0.61

^aDerived from DOE mid-term energy price forecasts and assumed to extend up to 10 years beyond mid-1995 to encompass a study period of 25 years beginning in 1980 (Ref. 1)

Appendix A

Block Diagram for LCOMP



* Δ MEANS THE DIFFERENTIAL COST BETWEEN SYSTEM 1 AND SYSTEM 2

Appendix B

LCOMP Listing

DSN*LCC(1) LCOMP

```

1      COMPILER (DATA=IBM)
2      DIMENSION P(10,5,3),G(10,5,3,3),BASEP(5),TOTCST(5,2),S(30,2)
3      INTEGER YR1,YR2
4      DIMENSION UNIT(5,3),ETYPE(10,3),SYS1(4),SYS2(4)
5      DIMENSION SECT(3),FUEL(30,2),OTH CST(30,2)
6      DATA ETYPE(1,1)/'ELECTR','ICITY','NATURA','L GAS','DISTIL','LATE',
7      1 'LIQUID',' GAS',2*'BLANK','ELECTR','ICITY','NATURA','L GAS',
8      2 'DISTIL','LATE','RESIDU','AL',2*'BLANK','ELECTR','ICITY',
9      3 'NATURA','L GAS','DISTIL','LATE','LIQUID',' GAS','COAL'/
10     DATA UNIT(1,1)/'KWH','CU FT ','GAL','GAL','BLANK',
11     1 'KWH','CU FT ','GAL','GAL','BLANK',
12     2 'KWH','CU FT ','GAL','GAL','TON'/
13     REAL NREC,NRECT,INVCST,LCC,NS
14     INTEGER YRREC,YRREP
15     DIMENSION TOTAL(2),LCC(2),NRECT(2),TITLE(12),NNREC(2)
16     DIMENSION CONS(5,2),YRCHG(20,5,2),C(20,5,2),R(2)
17     DIMENSION NREC(10,2),YRREC(10,2),YRREP(10,2),FINSVG(2)
18     DIMENSION INVCST(3,2),RPCST(10,2),RPSVG(10,2),NREP(2)
19     DATA TITLE(1)/12*6H      /
20     DATA SYS1(1)/4*6H      /,SYS2(1)/4*6H      /
21     DATA SECT(1)/'RESID ','COMMER','INDUST'/
22     C
23     C FNP IS SINGLE PRESENT WORTH FACTOR
24     C   DEFINE FNP(D,IY)=1 /(1 +D)**IY
25     C UPW IS UNIFORM PRESENT WORTH FACTOR
26     C   DEFINE UPW(D,IY)=(1 -1 /(1 +D)**IY)/D
27     C
28     C   NAMELIST/LCCIN/IYEAR,NSTUDY,CONS,YRCHG,C,IEXIST,ISECT,
29     1 IREG,INVCST,R,NNREC,NREC,YRREC,RPCST,YRREP,RPSVG,NREP,
30     2 FINSVG,DISC,TITLE,SYS1,SYS2,IBASE
31     C
32     C
33     C G(I,J,K,L) IS ESCALATION RATE,I=DOE REGION,J=FUEL TYPE,
34     C   K=ECONOMIC SECTOR,L=ESCALATION PERIOD
35     C P IS 1980 FUEL PRICE ARRAY
36     C
37     C   COMMON/BLKDAT/G,P
38     C
39     C   DISC=10
40     C   ICASE=1
41     C   IC=1
42     C   READ(5,LCCIN)
43     C   XDISC=DISC
44     C   DISC=DISC/100
45     C   WRITE(6,100) (TITLE(I),I=1,12),IYEAR,NSTUDY,ISECT,SECT(ISECT),IREG
46     1 ,XDISC
47     7 IF (IC EQ 1) WRITE(6,103) IC,SYS1,INVCST(1,IC),R(IC)
48     IF (IC EQ 2) WRITE(6,103) IC,SYS2,INVCST(1,IC),R(IC)
49     100 FORMAT(///19X,'**** LIFE CYCLE COST ANALYSIS ****',//12A6//
50     A 80('*')//,T30,'**** INPUT VALUES ****',//,
51     1 2X,'* PROJECT DESCRIPTION *'//8X,'BASE YEAR=',I5/8X,'STUDY ',
52     2 'PERIOD=',I3,' YEARS'/8X,'FINANCIAL SECTOR=',I2,2X,A6/8X,
53     3 'DOE REGION NO =' ,I2/,8X,'DISCOUNT RATE=',F4.1,' PERCENT'//,

```

```

54      4 //,2X,'* NON-FUEL COSTS IN BASE YEAR $ *')
55 103  FORMAT(/5X'SYSTEM ',I1,' (',4A6,')',//,
56      4 8X'INITIAL INVESTMENT COST=',T60,F9 0/8X'ANNUALLY RECURRING

57      5 'O&M COST=',T60,F9 0)
58      IF (NNREC(IC) EQ 0) GO TO 1
59      N=NNREC(IC)
60      WRITE(6,101) (YRREC(I,IC),NREC(I,IC),I=1,N)
61 101  FORMAT(8X'NON-ANNUALLY-RECURRING O&M COSTS IN YEAR',
62      1 I3,' =',T60,F9 0)
63      1 IF (NREP(IC) EQ 0) GO TO 6
64      N=NREP(IC)
65      WRITE(6,102) (YRREP(I,IC),RPCST(I,IC),RPSVG(I,IC),I=1,N)
66      6 IC=IC+1
67      IF (IEXIST EQ 1 AND IC EQ 2) GO TO 7
68      WRITE(6,104)
69  C
70  C  DETERMINE NO YEARS DIFFERENCE BETWEEN START OF STUDY AND BASE
71  C  ENERGY PRICE TABLES
72  C
73      IDELYR=IYEAR-IBASE
74      DO 2 J=1,5
75      2 BASEP(J)=P(IREG,J,Isect)
76      IF (IDELYR EQ 0) GO TO 4
77      IPER=1
78  C
79  C  UPDATE BASE ENERGY PRICES TO COINCIDE WITH START OF STUDY
80  C
81      DO 3 I=1,IDELYR
82      IF (I GT 5) IPER=2
83      IF (I GT 10) IPER=3
84      DO 3 J=1,5
85      3 BASEP(J)=BASEP(J)*(G(IREG,J,Isect,IPER)/100 +1 )
86      4 WRITE(6,105) ICASE
87      DO 20 J=1,5
88      ESC=BASEP(J)
89      J0=J*2-1
90      J1=J0+1
91      YR1=1
92      YR2=YRCHG(1,J,ICASE)-1
93      IF (YR2 EQ -1) YR2=NSTUDY
94      IF (CONS(J,ICASE) GT 1 E-5) WRITE(6,106) ETYPE(J0,Isect),
95      1 ETYPE(J1,Isect),CONS(J,ICASE),UNIT(J,Isect),YR1,YR2
96      IPER=1
97      ICNT=1
98      DO 10 I=1,NSTUDY
99      IF (I NE YRCHG(ICNT,J,ICASE)) GO TO 5
100     CONS(J,ICASE)=C(ICNT,J,ICASE)
101     YR1=I
102     ICP=ICNT+1
103     YR2=YRCHG(ICP,J,ICASE)-1
104     IF (YR2 EQ -1) YR2=NSTUDY
105     WRITE(6,106) ETYPE(J0,Isect),ETYPE(J1,Isect),CONS(J,ICASE),
106     1 UNIT(J,Isect),YR1,YR2
107     ICNT=ICNT+1
108     5 IF (I+IDELYR GT 5) IPER=2
109     IF (I+IDELYR GT 10) IPER=3
110  C
111  C  ESC IS ENERGY PRICE EACH YEAR EXCLUDING INFLATION
112  C

```

```

113             ESC=ESC*(G(IREG,J,ISECT,1PER)/100 +1 )
114 C
115 C   TOTCST IS TOTAL COST OF EACH ENERGY TYPE FOR ENTIRE STUDY PERIOD
116 C   FUEL IS ANNUAL FUEL COST DISCOUNTED TO BASE $
117 C   DISCOUNTED TO PRESENT $
118 C
119             X=CONS(J,ICASE)*ESC*FNP(DISC,I)
120             FUEL(I,ICASE)=FUEL(I,ICASE)+X
121             TOTCST(J,ICASE)=TOTCST(J,ICASE)+X
122 C
123 C   S IS USED TO CALCULATE DIFFERENTIAL SAVINGS FOR SIMPLE PAYBACK
124             S(I,ICASE)=S(I,ICASE)+CONS(J,ICASE)*BASEP(J)
125 10 CONTINUE
126 C
127 C   TOTAL IS TOTAL ENERGY COST FOR EACH SYSTEM (PRESENT VALUE)
128             TOTAL(ICASE)=TOTAL(ICASE)+TOTCST(J,ICASE)
129 20 CONTINUE
130             KCNT=1
131             ICNT=1
132 C
133 C   CALCULATE PRESENT VALUE OF TOTAL NON-RECURRING O&M COSTS,
134 C   REPLACEMENT COSTS, AND SALVAGE COSTS
135 C
136 C   OTHCST IS ANNUAL SUM OF OTHER COSTS
137             OTHCST(1,ICASE)=INVCST(1,ICASE)
138             DO 31 I=1,NSTUDY
139             OTHCST(I,ICASE)=OTHCST(I,ICASE)+R(ICASE)*FNP(DISC,I)
140             S(I,ICASE)=S(I,ICASE)+R(ICASE)
141             IF (I NE YRREC(ICNT,ICASE)) GO TO 30
142             S(I,ICASE)=S(I,ICASE)+NREC(ICNT,ICASE)
143             OTHCST(I,ICASE)=OTHCST(I,ICASE)+NREC(ICNT,ICASE)*FNP(DISC,I)
144             NRECT(ICASE)=NRECT(ICASE)+NREC(ICNT,ICASE)*FNP(DISC,I)
145             ICNT=ICNT+1
146 30 IF (I NE YRREP(KCNT,ICASE)) GO TO 31
147             S(I,ICASE)=S(I,ICASE)+RPCST(KCNT,ICASE)-RPSVG(KCNT,ICASE)
148             INVCST(2,ICASE)=INVCST(2,ICASE)+RPCST(KCNT,ICASE)*FNP(DISC,I)
149             INVCST(3,ICASE)=INVCST(3,ICASE)+RPSVG(KCNT,ICASE)*FNP(DISC,I)
150             OTHCST(I,ICASE)=OTHCST(I,ICASE)+RPCST(KCNT,ICASE)*FNP(DISC,I)
151 1 -RPSVG(KCNT,ICASE)*FNP(DISC,I)
152             KCNT=KCNT+1
153 31 CONTINUE
154             INVCST(3,ICASE)=INVCST(3,ICASE)+FINSVG(ICASE)*FNP(DISC,NSTUDY)
155             S(NSTUDY,ICASE)=S(NSTUDY,ICASE)-FINSVG(ICASE)
156             OTHCST(I,ICASE)=OTHCST(I,ICASE)-FINSVG(ICASE)*FNP(DISC,I)
157 C
158 C
159 C   NOW R=PRESENT VALUE OF ANNUALLY RECURRING O&M COSTS
160 C
161             R(ICASE)=R(ICASE)*UPW(DISC,NSTUDY)
162 C
163 C   LIFE CYCLE COST
164 C
165             LCC(ICASE)=TOTAL(ICASE)+INVCST(1,ICASE)+R(ICASE)
166 1 +NRECT(ICASE)+INVCST(2,ICASE)-INVCST(3,ICASE)
167             ICASE=ICASE+1
168             IF (ICASE EQ 2 AND IEXIST EQ 1) GO TO 4
169             IF (IEXIST EQ 0) GO TO 41
170 C

```

```

171 C SAVINGS-TO-INVESTMENT RATIO
172 C
173     SIR=(TOTAL(2)-TOTAL(1)+R(2)-R(1)+NRECT(2)-NRECT(1))/
174     1 (INVCST(1,1)-INVCST(1,2)+INVCST(2,1)-INVCST(2,2)+INVCST(3,2)
175     2 -INVCST(3,1))
176 C
177 C NET SAVINGS
178 C
179     NS=LCC(2)-LCC(1)
180 C
181 C PAYBACK PERIOD
182 C
183     CONST=INVCST(1,1)-INVCST(1,2)
184     DO 40 I=1,NSTUDY
185     SAVE=SAVE+S(I,2)-S(I,1)
186     IF (SAVE GT CONST) GO TO 41
187 40 CONTINUE
188 C
189 41 IPB= I
190     WRITE(6,108)
191     IX=IEXIST+1
192     DO 50 I=1,IX
193     WRITE(6,105) I
194     DO 49 J=1,5
195     J0=J*2-1
196     J1=J0+1
197     IF (TOTCST(J,I) GT 1.E-5) WRITE(6,109) ETYPE(J0,ISECT),
198 1 ETYPE(J1,ISECT),TOTCST(J,I)
199 49 CONTINUE
200     WRITE(6,110) TOTAL(I),R(I),NRECT(I),INVCST(2,I),INVCST(3,I),LCC(I)
201 50 CONTINUE
202     IF (IEXIST EQ 0) STOP
203     WRITE(6,111) LCC(1),LCC(2),NS,SIR,IPB
204     WRITE(6,112)
205     DUM=0
206     CUM=INVCST(1,2)-INVCST(1,1)
207     WRITE(6,113) I0,DUM,INVCST(1,1),INVCST(1,1),DUM,
208 1 INVCST(1,2),INVCST(1,2),CUM
209     DO 60 I=1,NSTUDY
210     CUM1=CUM1+FUEL(I,1)+OTH CST(I,1)
211     CUM2=CUM2+FUEL(I,2)+OTH CST(I,2)
212     CUM=CUM2-CUM1
213     WRITE(6,113) I,FUEL(I,1),OTH CST(I,1),CUM1,FUEL(I,2),OTH CST(I,2)
214 1 ,CUM2,CUM
215 60 CONTINUE
216 102 FORMAT(8X'REPLACEMENT OCCURS IN YEAR',I3,' '/12X,
217 1 'REPLACEMENT COST=',T60,F9 0/12X,'SALVAGE VALUE=',
218 2 T60,F9 0)
219 104 FORMAT(/80('*'),//T20,'**** SUMMARY OF ANNUAL FUEL CONSUMPTION
220 1 '****',//)
221 105 FORMAT(/5X,'SYSTEM ',I1)
222 106 FORMAT(8X,2A6,' --',F10 0,A6,'/YR (DURING YEAR',I3,'-',I3,'')
223 108 FORMAT(/80('*'),//T25,'**** ANALYSIS RESULTS ****'//,T24,
224 1 '(PRESENT VALUES IN BASE YEAR $)')
225 109 FORMAT(8X'TOTAL ',2A6,' COSTS=',T60,F9 0)
226 110 FORMAT(8X'TOTAL, ALL ENERGY COSTS=',T59,F10 0/
227 1 8X'TOTAL ANNUALLY RECURRING O & M COSTS=',T59,F10 0/
228 2 8X'TOTAL NON-RECURRING O & M COSTS=',T59,F10 0/
229 3 8X'TOTAL REPLACEMENT COSTS=',T59,F10 0/

```

```

230      4 8X'TOTAL SALVAGE VALUES=',T59,F10 0/
231      5 8X'LIFE CYCLE COST=',T59,F10 0)
232      111 FORMAT(/2X'COMPARISON RESULTS '/8X'LIFE CYCLE COST, SYSTEM 1='
233      1 ,T59,F10 0/8X,'LIFE CYCLE COST, SYSTEM 2=',T59,F10 0,/
234      2 8X,'NET SAVINGS=',T59,F10 0/
235      1 8X'SAVINGS-T0-INVESTMENT RATIO=',T63,F6 3/8X
236      2 'PAYBACK DURING YEAR ',I2,' ,BASED ON SIMPLE PAYBACK ANALYSIS'
237      3 //80('*'))
238      112 FORMAT(/,'1',T25,'**** CASH FLOW ANALYSIS ****'//
239      1 T22,'(PRESENT VALUE IN BASE YEAR $)',///,T19,'SYSTEM 1',
240      1 T40,'*',T53,'SYSTEM 2',T74,'*',T79,'SYSTEM 1',/2X,'YR',T9,'FUEL',
241      2 T18,'NET OTHER',T31,'CUMUL ',T40,'*',T45,'FUEL',T54,'NET OTHER',
242      3 T66,'CUMUL ',T75,'*',T78,'VS SYSTEM2',/T9,'COST',T20,'COSTS',
243      4 T29,'TOTAL COST',T40,'*',T45,'COST',T56,'COSTS',T64,'TOTAL COST',
244      5 T75,'*',T77,'CUMUL SAVINGS'/T2,37('-'),' * ',32('-'),' * ',
245      6 15('-'))
246      113 FORMAT(I4,T6,F10 0,T17,F9 0,T29,F9 0,T40,'*',T42,F9 0,T53,
247      1 F9 0,T64,F9 0,T75,'*',T79,F9 0)
248      STOP
249      END

```

DSN*LCC(1) BLK2

```

1      COMPILER (DATA=IBM)
2      C ENERGY PRICES AND ESCALATION RATES AS PUBLISHED IN
3      C NBS HANDBOOK 135 (DEC 1980)
4      BLOCK DATA
5      COMMON/BLKDAT/G(10,5,3,3),P(10,5,3)
6      DATA G(1,1,1,1)/4*- 02,- 01,3*- 02,- 01,- 02,
7      1 2*1 78,1 77,2*1 75,1 76,1 75,1 76,1 77,1 78,
8      2 3 38,3 39,3 38,3 39,3 37,2*3 38,3 39,3 37,3 37,
9      3 20*0 ,
10     4 - 01,- 02,2*- 01,- 02,2*- 01,- 02,- 01,- 02,
11     5 1 77,1 76,1 75,1 73,2*1 74,1 73,1 75,2*1 76,
12     X 3 38,3 39,2*3 38,3 39,3 38,3 4,3*3 38,
13     X 7 53,7 52,7 52,7 53,7 55,7 52,7 55,7 51,7 5,7 57,10*0 ,
14     X - 01,- 03,- 01,- 02,- 01,- 02,- 02,- 01,- 01,- 05,
15     6 1 8,1 76,1 78,1 76,1 77,1 79,1 74,1 8,1 79,1 79,
16     7 3 39,3 4,3*3 39,3 38,3 4,3 37,2*3 38,
17     8 7 54,3*7 53,7 54,7 53,7 54,7 53,7 51,7 55,
18     9 9 62,9 51,9 63,9 58,9 49,9 62,9 5,9 3,2*9 56,
19     DATA G(1,1,1,2)/- 02,- 61, 87,1 72,1 04,1 53,- 59,- 2 73, 47,3 85,
20     1 3 33,2 74,3 15,2 38,2 84,4 53,3 43,3 95,1 44,3 86,
21     2 2 81,2 8,2 74,2 71,2 91,2 83,2 94,2 82,2 97,2 97,
22     3 20*0 ,
23     4 - 19,- 64, 89,1 67,1 07,1 62,- 63,- 2 96, 43,3 97,
24     5 3 88,3 2,3 6,2 82,3 15,5 26,3 85,4 22,1 66,4 5,
25     6 2 91,2 88,2 89,2 88,2 99,2 93,3 01,2 91,3 09,3 09,
26     7 2 66,2 68,2 52,2 71,2 7,2 71,2 69,2 67,2 84,2 92,
27     8 10*0 ,
28     9 - 23,- 98,1 19,2 2,1 47,2 03,- 78,- 4 05, 55,7 89,
29     1 3 81,4 46,8 89,8 82,6 98,5 88,11 74,4 95,3 26,6 2,
30     2 3 47,2 88,2*2 85,2 99,2 94,3 02,2 9,2*3 09,
31     3 3 53,2 6,2 58,2 73,2*2 71,2 7,2 69,2 87,2 86,
32     4 1 47,1 65,1 97,1 69,1 67,1 45,1 76,0 ,1 36,2 39/
33     DATA G(1,1,1,3)/- 3 55,- 42,- 36, 48,- 13,- 26,- 02,- 2 47,- 2 35,
34     1 1 1,1 92,1 68,1 56,1 01, 97,2 33,1 23,2 36, 26,- 3 09,
35     2 3 97,3 95,3 89,3 86,4 1,4 ,4 13,4 06,4 13,4 13,
36     3 20*0 ,
37     4 - 3 6,- 44,- 37, 48,- 14,- 28,- 03,- 2 7,- 2 21,1 13,
38     5 2 16,1 93,1 8,1 18,1 11,2 67,1 38,2 49, 29,- 3 58,

```

```

39      6 4 08,4 09,4 07,4 06,4 2,4 12,4 22,4 18,4 27,4 27,
40      7 4 47,4 41,4 26,4 5,2*4 56,4 54,4 6,4 68,4 82,
41      8 10*0 ,
42      9 -4 45,- 68,- 5, 63,- 19,- 34,- 03,-3 91,-2 78,1 97,
43      1 4 72,4 6,- 95,3 41,3 92,2 89,9 6,2 84, 93,3 99,
44      2 4 08,4 1,4 02,4 02,4 2,4 13,4 22,4 16,2*4 27,
45      3 4 43,4 35,4 28,4 52,4 58,4 57,4 53,4 64,4 68,4 71,
46      4 -3 47,1 ,1 12, 79, 98, 6, 77, 62, 94,-2 87/
47      DATA P(1,1,1)/ 091, 086, 064, 049, 059,3* 064, 069, 025,
48      1 005, 005,7* 004, 005,
49      2 997,1 008,1 035,1 044, 968, 999, 96, 976, 949, 949,
50      3 20*0 ,
51      4 089, 082, 063, 05, 057, 06, 061, 06, 073, 024,
52      5 3* 004,5* 003,2* 004,
53      6 963, 972, 978, 979, 94, 963, 933, 943, 908, 908,
54      7 742, 753, 792, 742, 744, 744, 748, 73, 708, 684,
55      8 10*0 ,
56      9 074, 054, 047, 037, 041, 048, 049, 045, 058,.011,

57      1 006, 006, 005,2* 004,3* 003, 006, 005,
58      2 936, 969, 992, 99, 94, 96, 932, 949, 908, 908,
59      3 717, 766, 781, 738, 741, 742, 745, 723, 707, 701,
60      4 41 175,36 675,31 95,35 775,29 025,33 525,25 875,20 25,
61      5 42 525,34 875/
62      END

```

Appendix C

LCOMP Inputs

THE INPUT DATA TO LCOMP IS VIA NAMELIST "\$LCCIN" WHICH IS READ ON I/O UNIT 5
THE NAMELIST VARIABLES ARE AS FOLLOWS

TITLE (I)	=	CASE TITLE, UP TO 48 CHARACTERS IN LENGTH ENTER AS TITLE (I) = "SOME TITLE "
IYEAR	=	CALENDAR YEAR WHEN STUDY BEGINS
NSTUDY	=	LENGTH OF STUDY PERIOD (YEARS).
IBASE	=	BASEYEAR (CALENDAR) FOR ENERGY PRICE AND ESCALATION TABLES IT SHOULD BE NOTED THAT ALL COSTS AND SALVAGE VALUES ARE IN BASE-YEAR DOLLARS
IEXIST	=	0 IF ONLY A NEW BUILDING, OR ONLY ONE SYSTEM IS TO BE CONSIDERED = 1 IF TWO SYSTEMS ARE BEING COMPARED
ISECT	=	1 IF THE ECONOMIC SECTOR IS RESIDENTIAL = 2 COMMERCIAL = 3 INDUSTRIAL.
IREG	=	DOE REGION
1	⇒	MAINE, NEW HAMPSHIRE, VERMONT, MASSACHUSETTS, CONNECTICUT, RHODE ISLAND
2	⇒	NEW YORK, NEW JERSEY, PUERTO RICO, VIRGIN ISLANDS
3	⇒	PENNSYLVANIA, MARYLAND, WEST VIRGINIA, VIRGINIA, WASHINGTON, D C., DELAWARE
4	⇒	KENTUCKY, TENNESSEE, NORTH CAROLINA, MISSISSIPPI, ALABAMA, GEORGIA, FLORIDA, CANAL ZONE
5	⇒	MINNESOTA, WISCONSIN, MICHIGAN, ILLINOIS, INDIANA, OHIO
6	⇒	TEXAS, NEW MEXICO, OKLAHOMA, ARKANSAS, LOUISIANA
7	⇒	KANSAS, MISSOURI, IOWA, NEBRASKA
8	⇒	MONTANA, NORTH DAKOTA, SOUTH DAKOTA, WYOMING, UTAH, COLORADO
9	⇒	CALIFORNIA, NEVADA, ARIZONA, HAWAII, PACIFIC ISLANDS, SAMOA, GUAM
10	⇒	WASHINGTON, OREGON, IDAHO, ALASKA
DISC	=	DISCOUNT RATE, PERCENT (i e , ENTER 7% AS 7 0)
CONS(J,K)	=	BASEYEAR CONSUMPTION OF THE J th ENERGY TYPE, WHERE J IS AN INTEGER WHICH TAKES THE VALUES
J = 1	⇒	ELECTRICITY (kW)
J = 2	⇒	NATURAL GAS (cu ft)
J = 3	⇒	DISTILLATE (gal)
J = 4	⇒	RESIDUAL (gal), COMMERCIAL AND INDUSTRIAL ONLY
J = 5	⇒	COAL (ton), INDUSTRIAL ONLY
K	⇒	SYSTEM BEING ANALYZED
K = 1	⇒	NEW, RETROFIT, OR SOLAR ENERGY SYSTEM
K = 2	⇒	EXISTING OR CONVENTIONAL SYSTEM

SUBSEQUENT REFERENCES TO ARRAY INDEX (K) INDICATE SYSTEM TYPE.

YRCHG(I,J,K)	=	STUDY YEAR (NOT CALENDAR YEAR) IN WHICH A CHANGE IN THE AMOUNT OF ANNUAL ENERGY CONSUMPTION OCCURS
J	⇒	ENERGY TYPE (AS ABOVE)
K	⇒	SYSTEM (AS ABOVE)
I	⇒	I th CHANGE FOR A PARTICULAR SYSTEM AND ENERGY TYPE
C(I,J,K)	=	NEW ENERGY CONSUMPTION AMOUNT CORRESPONDING TO YRCHG (I,J,K). IF AMOUNTS OF ANNUAL ENERGY CONSUMPTION DO NOT CHANGE, YRCHG AND C MAY BE OMITTED FROM THE NAMELIST INPUT
INVCST(I,K)	=	INITIAL INVESTMENT COST FOR SYSTEM K (IF EXISTING BUILDING, INPUT CURRENT SALVAGE VALUE)
R(K)	=	ANNUALLY RECURRING O&M COST FOR SYSTEM (K).
NREC(I,K)	=	I th NON-RECURRING O&M COST FOR SYSTEM (K)
YRREC(I,K)	=	YEAR WHEN NON-RECURRING O&M COST OCCURS, CORRESPONDING TO NREC (I,K)
NNREC(K)	=	THE NUMBER OF NON-RECURRING O&M COSTS FOR SYSTEM (K)
RPCST(I,K)	=	I th REPLACEMENT COST FOR SYSTEM (K)
YRREP(I,K)	=	YEAR WHEN REPLACEMENT OCCURS, CORRESPONDING TO RPCST (I,K).
RPSVG(I,K)	=	SALVAGE VALUE OF EQUIPMENT BEING REPLACED
NREP(K)	=	THE NUMBER OF NON-RECURRING O&M COSTS FOR SYSTEM (K)

FINSVG(K) = FINAL SALVAGE VALUE OF SYSTEM (K).
 SYS1 (I) = SYSTEM (1) TITLE, UP TO 24 CHARACTERS IN LENGTH, SHOULD BE NAMES
 LIKE NEW SYSTEM, RETROFIT SYSTEM, SOLAR ENERGY SYSTEM
 SYS2 (I) = SYSTEM (2) TITLE, UP TO 24 CHARACTERS IN LENGTH, SHOULD BE NAMES
 LIKE EXISTING SYSTEM, CONVENTIONAL SYSTEM.

THE OTHER INPUT TO LCOMP CONSISTS OF ENERGY BASEYEAR PRICES AND ESCALATION
 RATES THIS DATA IS CONTAINED IN A BLOCKDATA SUBPROGRAM AS SHOWN IN THE PROGRAM
 LISTING IN APPENDIX A. THE BLOCKDATA IS NOT USUALLY CONSIDERED TO BE A CASE-
 DEPENDENT INPUT AS IS THE NAMELIST THE BLOCKDATA VARIABLES ARE

G(I,J,K,L) = ENERGY PRICE ESCALATION RATES
 I ⇒ DOE REGION NUMBER AS DEFINED BY IREG
 J ⇒ ENERGY TYPE AS DEFINED BY CONS
 K ⇒ ECONOMIC SECTOR AS DEFINED BY ISECT
 L ⇒ ESCALATION PERIOD
 L = 1 ⇒ 1980-1984
 L = 2 ⇒ 1985-1989
 L = 3 ⇒ 1990 AND BEYOND } FOR DATA IN APPENDIX B
 P(I,J,K) = BASELINE ENERGY PRICES (1980 IS BASELINE YEAR FOR THE DATA SHOWN IN
 APPENDIX B.)

IT SHOULD BE NOTED THAT THE STRUCTURE OF THIS DATA CAN BE CHANGED BY THE USER,
 PROVIDING A MEANS FOR THE USER TO DEFINE THEIR OWN ENERGY PRICES, ESCALATION
 RATES, ESCALATION PERIODS, ETC

Appendix D

LCOMP Output for a Sample Problem

```

      **** LIFE CYCLE COST ANALYSIS ****

      NATIONAL ADMINISTRATION, AUTOMATIC CONTROL SYSTEM

      *****

      **** INPUT VALUES ****

      * PROJECT DESCRIPTION *

      BASE YEAR= 1980
      STUDY PERIOD= 30 YEARS
      FINANCIAL SECTOR= 2 (COMMER)
      DOE REGION NO = 3
      DISCOUNT RATE= 7 0 PERCENT

      * NON-FUEL COSTS IN BASE YEAR $ *

      SYSTEM 1 (RETROFIT SYSTEM      )

      INITIAL INVESTMENT COST=                1500000
      ANNUALLY RECURRING O&M COST=            145000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 10 =      40000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 20 =      40000

      SYSTEM 2 (EXISTING SYSTEM      )

      INITIAL INVESTMENT COST=                260000
      ANNUALLY RECURRING O&M COST=            120000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 5 =        25000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 10 =        25000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 15 =        25000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 18 =        60000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 20 =        25000
      NON-ANNUALLY-RECURRING O&M COSTS IN YEAR 25 =        25000
      REPLACEMENT OCCURS IN YEAR 10
      REPLACEMENT COST=                        150000
      SALVAGE VALUE=                          25000
      REPLACEMENT OCCURS IN YEAR 20
      REPLACEMENT COST=                        150000
      SALVAGE VALUE=                          25000

      *****
  
```

Fig D-1. Input values

**** SUMMARY OF ANNUAL FUEL CONSUMPTION ****

SYSTEM 1

ELECTRICITY	--	9190000	KWH	/YR (DURING YEAR 1- 4)
ELECTRICITY	--	9080000	KWH	/YR (DURING YEAR 5- 9)
ELECTRICITY	--	8790000	KWH	/YR (DURING YEAR 10- 30)
NATURAL GAS	--	386800000	CU FT	/YR (DURING YEAR 1- 4)
NATURAL GAS	--	369000000	CU FT	/YR (DURING YEAR 5- 9)
NATURAL GAS	--	362400000	CU FT	/YR (DURING YEAR 10- 30)

SYSTEM 2

ELECTRICITY	--	9982200	KWH	/YR (DURING YEAR 1- 30)
NATURAL GAS	--	433600000	CU FT	/YR (DURING YEAR 1- 30)

Fig D-2. Summary of annual fuel consumption

**** CASH FLOW ANALYSIS ****

(PRESENT VALUE IN BASE YEAR \$)

YR	FUEL COST	SYSTEM 1 NET OTHER COSTS	CUMUL TOTAL COST	*	FUEL COST	SYSTEM 2 NET OTHER COSTS	CUMUL TOTAL COST	*	SYSTEM 1 VS SYSTEM2 CUMUL SAVINGS
0	0	1500000	1500000	*	0	260000	260000	*	-1240000
1	2012325	1635514	3647839	*	2236979	372150	2609129	*	-1038711
2	1904690	126649	5679178	*	2117554	104813	4831496	*	-847683
3	1802920	118363	7600461	*	2004622	97956	6934074	*	-666387
4	1706687	110620	9417768	*	1897823	91547	8923444	*	-494323
5	1555382	103383	11076533	*	1796818	103383	10823645	*	-252888
6	1495634	96620	12668787	*	1728372	79961	12631978	*	-36809
7	1438374	90299	14197460	*	1662749	74730	14369457	*	171997
8	1383490	84391	15665341	*	1599823	69841	16039121	*	373780
9	1330873	78870	17075084	*	1539472	65272	17643865	*	568781
10	1253252	94045	18422380	*	1481583	137254	19262702	*	840322
11	1186382	68888	19677650	*	1402807	57011	20722520	*	1044870
12	1123172	64382	20865203	*	1328326	53281	22104127	*	1238924
13	1063416	60170	21988789	*	1257899	49796	23411822	*	1423033
14	1006921	56234	23051943	*	1191299	46538	24649659	*	1597716
15	953503	52555	24058001	*	1128313	52555	25830526	*	1772525
16	902990	49117	25010107	*	1068738	40648	26939912	*	1929805
17	855219	45903	25911229	*	1012385	37989	27990286	*	2079057
18	810038	42900	26764167	*	959075	53256	29002616	*	2238449
19	767301	40094	27571562	*	908638	33181	29944435	*	2372873
20	726874	47808	28346243	*	860917	69773	30875125	*	2528881
21	688627	35019	29069890	*	815760	28982	31719866	*	2649976
22	652441	32728	29755059	*	773026	27086	32519977	*	2764918
23	618200	30587	30403846	*	732581	25314	33277871	*	2874025
24	585798	28586	31018230	*	694301	23658	33995829	*	2977599
25	555133	26716	31600079	*	658065	26716	34680610	*	3080530
26	526110	24968	32151157	*	623762	20663	35325034	*	3173877
27	498638	23335	32673130	*	591286	19312	35935631	*	3262502
28	472632	21808	33167570	*	560537	18048	36514216	*	3346646
29	448012	20382	33635963	*	531421	16868	37062504	*	3426541
30	424702	-30214	34030451	*	503849	12480	37578833	*	3548382

Fig D-3 Cash flow analysis

*** ANALYSIS RESULTS ***

(PRESENT VALUES IN BASE YEAR \$)

SYSTEM 1

TOTAL ELECTRICITY COSTS=	7095018
TOTAL NATURAL GAS COSTS=	23654717
TOTAL, ALL ENERGY COSTS=	30749735
TOTAL ANNUALLY RECURRING O & M COSTS=	1799311
TOTAL NON-RECURRING O & M COSTS=	30671
TOTAL REPLACEMENT COSTS=	0
TOTAL SALVAGE VALUES=	49263
LIFE CYCLE COST=	34030454

SYSTEM 2

TOTAL ELECTRICITY COSTS=	7894460
TOTAL NATURAL GAS COSTS=	27774317
TOTAL, ALL ENERGY COSTS=	35668777
TOTAL ANNUALLY RECURRING O & M COSTS=	1489085
TOTAL NON-RECURRING O & M COSTS=	68413
TOTAL REPLACEMENT COSTS=	115015
TOTAL SALVAGE VALUES=	22453
LIFE CYCLE COST=	37578836

COMPARISON RESULTS

LIFE CYCLE COST, SYSTEM 1=	34030454
LIFE CYCLE COST, SYSTEM 2=	37578836
NET SAVINGS=	3548382
SAVINGS-TO-INVESTMENT RATIO=	4.231
PAYBACK DURING YEAR 6, BASED ON SIMPLE PAYBACK ANALYSIS	

Fig D-4 Analysis results

GCF-NOCC Reconfiguration

J P McClure
DSN Data Systems Section

The equipment and computer programs in the Network Operations Control Center (NOCC) and the Ground Communications Facility's (GCF's) Central Communication Terminal at JPL have been rearranged and supplemented to provide an improved operational capability

The computer portion of the GCF's Central Communication Terminal (CCT) in the basement of the Space Flight Operations Facility was designed and implemented separate from the Network Data Processing (NDP) portion of the Network Operations Control Center (NOCC). These independent designs, implemented at different times, used more computers and programs than an integrated design would have required. The desire to reduce the computer count and to provide an improved operational capability led to the GCF-NOCC reconfiguration effort.

This effort, which was completed in April 1981

- (1) Requires fewer minicomputers (16 vs 20)
- (2) Requires fewer computer programs (6 vs 8)
- (3) Significantly reduces magnetic tape handling
- (4) Provides a much improved monitor and control capability
- (5) Reduces operator requirements and provides the base for a two-operator CCT

The reconfigured GCF-NOCC is shown in Fig 1. Digital communications circuits from the Deep Space Communica-

tions Complexes (DSCCs) are routed through circuit switches to the transmission terminal equipment for each circuit. The terminal equipment includes error detection encoding-decoding hardware and distribution amplifiers to directly route wideband data to the MCCC and the DSN's VLBI correlation processors. All DSCC data are contained in standardized blocks which include an address at the beginning of the block plus an error detection code at the end of the block.

Data then flow to the Error Correction and Switching (ECS) processors for routing. There are four ECS processors. Three on-line units share the load, with one hot backup which can be switched on line in a minute or less to replace a failed processor. The ECS processors route the data blocks

- (1) To the addressee named in the block (usually the MCCC)
- (2) To the Data Records Generation processors if a permanent Intermediate Data Record (IDR) is to be produced
- (3) To the Network Communications Equipment (NCE)
- (4) Optionally, to a Front End Record (FER) magnetic tape unit driven by the ECS itself

The Network Communications Equipment (NCE) consists of two processors, a prime and backup, which route blocks to the proper NOCC computers (RTMs and Support). The NCE also routes display data generated by the RTMs to the Digital Display Processor (DDP) which in turn forwards the formatted information to the Video Display Processor (VAP). The VAP drives the network display devices which depict the status of the entire DSN. The DDP and VAP assemblies are each dual units, prime and backup.

The Data Records Generation assembly consists of three on-line and one hot backup unit. Real-time data received from the ECS are subdivided into streams according to spacecraft and data type and then recorded. High-speed streams (< 3-4 kbps) are written on disc while wideband streams are recorded on magnetic tape. Gaps in the data are detected as the real-time stream is recorded. After the tracking pass is complete,

the missing data are recalled from the DSCC, the gaps in the record are filled and a tape-recorded Intermediate Data Record (IDR) is written.

The Central Communications Monitor (CCM) receives status reports from all of the GCF's processors, including the Communications Monitor and Formatter (CMF) machines at the DSSs. The status information is correlated, formatted and presented to Communications Operations personnel. The CCM is also a central control point for all of the GCF computers in the Central Communications Terminal and for the hardware switches.

The reconfigured GCF-NOCC capability was completed in April 1981. The Voyager 2 encounter with Saturn was supported by this new capability which performed as designed.

References

- 1 McClure, J. P., "GCF-NOCC Reconfiguration," in *The Deep Space Network Progress Report 42-55*, pp. 86-89, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 1980.
- 2 Bremner, D. S., and Hung, C. K., "Ground Communications Facility and Network Operations Control Center Reconfiguration," in *The Deep Space Network Progress Report 42-58*, pp. 108-109, Jet Propulsion Laboratory, Pasadena, Calif., July 1980.

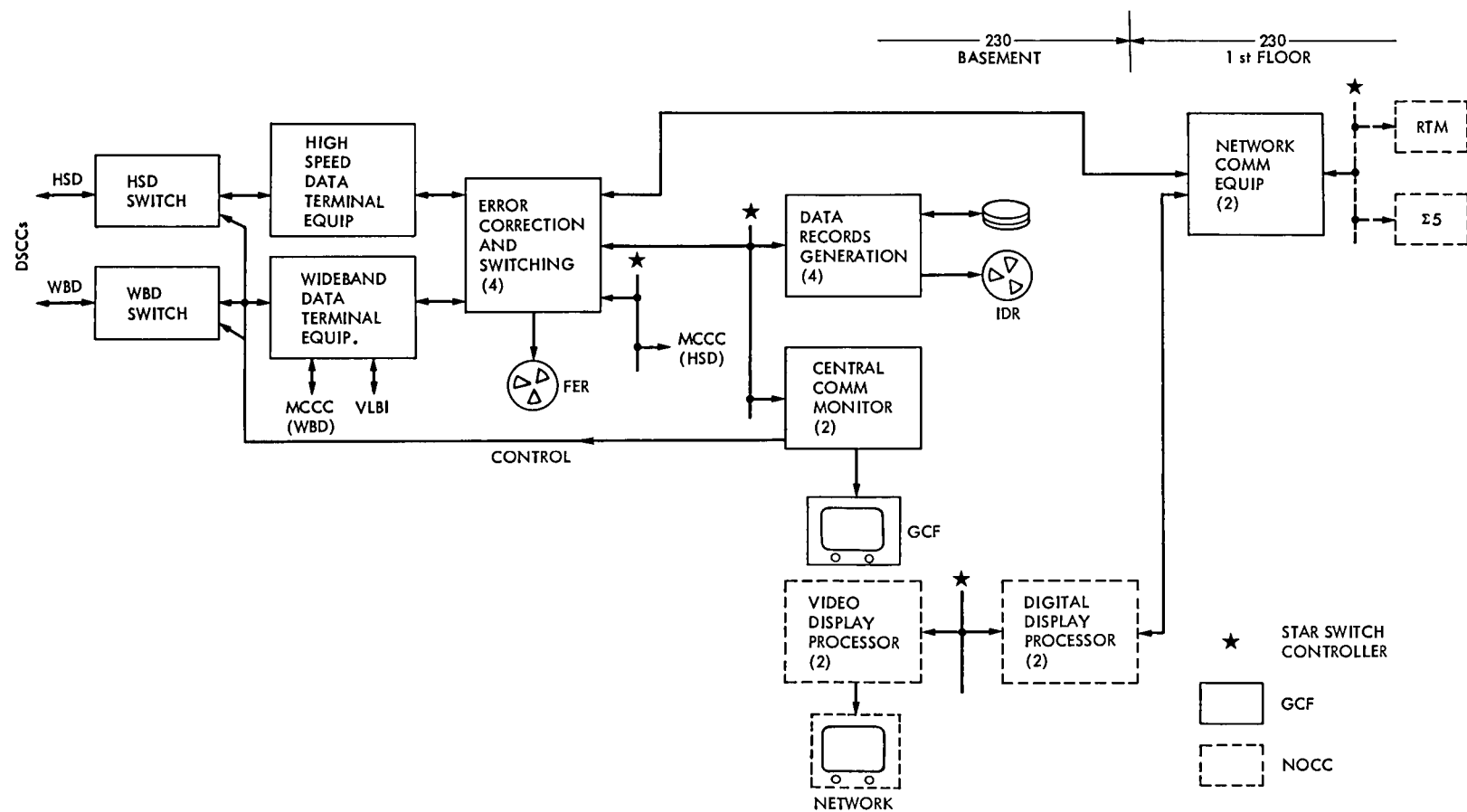


Fig 1. Reconfigured GCF-NOCC

Antenna Azimuth Bearing Model Experiment

H McGinness
DSN Engineering Section

The concluding description of the test is presented and reasons given for expecting a long service life from a large-scale prototype.

I. Introduction

In *DSN Progress Report 42-53* (Ref 1), a partial report was given on the performance of a reduced-scale model of an antenna azimuth bearing. This bearing is of the wheel and track type and employs a novel suspension of the wheel. The following article completes the description of the model tests and discusses various aspects of the design. It is necessary that the reader refer to the partial report (Ref 1) in order to understand this concluding part.

II. Model Running Tests

For the first two series of tests the model turntable was supported by three pairs of wheels as shown in Figs 1 and 2 of Ref 1. The load per wheel was 5560 N and produced a nominal hertz stress of $86,184 (10^4) \text{ N/m}^2$, based upon the net width of the contact area, which is the wheel width minus twice the corner radius.

The first series of tests was run for 40 60 hours, half of which was in the clockwise direction. No lubricant was added to the track, but a reddish brown color developed on top of the wear strips over the portion contacted by the wheels. The same color developed on the wheel rolling surfaces. It was decided to regrind the wear strip upper surfaces and the wheel rolling surfaces before starting the second series of tests. Upon

removal of the wear strips it was seen that there was fretting corrosion at the interface of the wear strips and their support ring. It was then obvious that fretting products had migrated through the wear strip mitered joints and accounted for the discoloration of the wheels and top surfaces of the wear strips. A photograph of a typical corroded wear strip and support ring section is shown in Fig 1.

There was no evidence of spalling or other fatigue failure on either the wheels or wear strips, however, there were minute pits probably caused by small broken particles from the edges of the mitered joints. (Later it will be shown that this will not occur on the prototype.) Figures 2 and 3 are, respectively, photographs of a wheel surface and wear strip after 40 6 hours of running.

Before regrinding the wheels, their diameters were measured and it was estimated that their diameter reductions through wear did not exceed 0 008 mm. Approximately 0 050 mm of material was removed from the top surfaces of the wear strips and from the rolling surfaces of the wheels by regrinding. Before reinstalling the wear strips, various lubricants were put between them and the support ring. These are shown in Fig 4.

The second series of tests was run for a total of 81 54 hours, approximately half of which was in the clockwise

direction. The total running time for the first two series of tests was 122 14 hours or 45869 revolutions of the turntable. There was no evidence of fatigue failure of the wheels or wear strips. During the second series some of the lubricants placed on the bottom surfaces of the wear strips were extruded through the wear strip mitered joints and smeared over the upper surfaces, thus producing an inadvertent lubrication of the wheel track surface. After 31 66 hours of the second series of tests, a track lubricator was installed on each end of each pair of wheels. The lubricator consisted of a cylindrical rod of oil-impregnated plastic (SKF POLYOIL) approximately 11 mm in diameter, one end of which was lightly spring-loaded against the wear strip. It is claimed that the very thin film of oil deposited soon hardens to a waxlike substance which serves as a wheel-track lubricator yet does not collect foreign objects as would a conventional oil. Such a device serves also to brush large foreign objects away from the wheel path.

After 49 13 hours of running the second series of tests, 10 of the wear strips were removed and examined for fretting. Only two of the several lubricants had successfully prevented fretting. The results are listed in Table 1. The fretted wear strips were cleaned, coated with Silver Goop, replaced, and the second series of tests was resumed and run for 32 41 hours more. All wear strips were then removed and examined. Qualitative results are listed in Table 2.

For the third series of tests the wear strips were cleaned by regrounding slightly on both top and bottom surfaces. The bottom surfaces of some were plated as indicated in Fig 5, whereas the remaining seven had the proprietary product CORTEC VCJ-309, a white powdered substance, put at the interface with the support ring.

Each wheel assembly was modified by replacing each used pair by an unused single wheel.¹ The three wheels were then adjusted so that their tangential misalignment to the track was not more than 30 arcseconds. The ballast weights were then adjusted so that the reaction on each of the two opposite wheels was 8749 N and the reaction on the third wheel was 8913 N. These values include the small effect of centrifugal force when rotating at the speed of 375 55 turns per hour. These loads are approximately 1 60 times the wheel loads used during the first two series of tests. The value of 1 6 is obtained from Eq (6) of Ref 1 by setting the scale factor λ at 0 042. Thus when loaded to this value the model is of scale 0 042 from a wheel-track fatigue life standpoint.

The third test series was begun and run intermittently over five days for a total of 38 24 hours. Approximately half was in

the clockwise direction. The wheels and wear strips were visually inspected and appeared to be free from pits and spalling. There were some abrasion marks near the edge of each wheel but these were caused by the wheels hitting one of the wear strip holddown screw heads.

Five of the wear strips were removed and examined. Numbers 5, 16 and 18, which were plated respectively with copper, nickel, and tin, had lost most of the plating and fretting corrosion occurred. The chromium plating on number 2 was intact and clean. All the others, which had CORTEC at the interface, looked very good with only small traces of fretting.

For the fourth series of tests each wheel loading was reduced to 5674 N and each wheel was deliberately misaligned by 40 arcminutes with respect to a track tangent line. The turntable was run for a total of 19.28 hours, 11.16 hours of which was in the clockwise direction. An examination of the wheels and wear strips revealed a small amount of spalling at four places on wear strip number 15. This was the first and only evidence of fatigue failure. The wear strips had undergone a total of 179 66 hours. The test conditions are summarized in Table 3.

III. Drag Resistance Tests

In addition to the drag resistance tests reported in Ref. 1, the following were made. All the track lubricators were removed and the drive chain disconnected. The turntable was pushed with a force gage located at a radius of 1 84 m. When the turntable was pushed at a comfortable walking speed the gage read between 57 and 45 N. Then the wheel alignment was restored to the original good alignment of within 30 arcseconds of perfection, and push tests made. The force gage read between 40 and 31 N. The measured rolling resistance coefficient f_m is defined as

$$f_m = \frac{T}{WR} \quad (1)$$

where

T is the measured torque

W is the total weight on all wheels

R is the mean radius of the track

Using 17000 N for W and 1.77 m for R , the following are calculated

$$f_{m \text{ aligned}} = \frac{(1.84) \times \frac{40}{31}}{17000 (1.77)} = \frac{0.0024}{0.0019} \quad (2)$$

¹The drawings which show the detailed construction of the model are JPL drawings 9472187 through 9472198.

$$f_{m \text{ misaligned by } 40 \text{ arcmin}} = \frac{(1.84) \times \frac{57}{45}}{17000 (1.77)} = \frac{0.0035}{0.0027} \quad (3)$$

From Eq (8) of Ref 1, the following corresponding theoretical coefficient f_T may be obtained

$$f_T = f_1 + f_2 \frac{d}{D} + \mu \sin \theta \quad (4)$$

where

f_1 is the factor for roller thrust bearings

f_2 is the factor for radial ball bearings on the wheel axis

μ is the coefficient of sliding friction

d is the wheel bearing bore diameter

D is the wheel diameter

Using the following values $f_1 = 0.0011$, $f_2 = 0.0015$, $d/D = 0.30$, $\mu = 0.20$, $\sin 40 \text{ arcmin} = 0.0116$, Eq (4) yields

$$f_{T \text{ aligned}} = 0.0011 + 0.0015 (0.30) + 0.20 (0.00015) = 0.0016 \quad (5)$$

$$f_{T \text{ misaligned by } 40 \text{ arcmin}} = 0.0011 + 0.0015 (0.30) + 0.20 (0.0116) = 0.0039 \quad (6)$$

The results of (5) and (6) are to be compared to the results of (2) and (3) respectively. The comparison suggests that Eq (4) applies reasonably well to a large bearing of this type.

IV. Static Tests

In order to demonstrate that the wheel suspension system would safely support considerably more than its nominal load, it was overloaded under the following conditions. All of the ballast weights were stacked together at the position shown on the vertical centerline of Fig 1 of Ref 1. There were a total of three properly aligned wheels supporting the turntable. The total ballast weight was 27376 N and the reaction of the wheel nearest the ballast was 19547 N. The turntable was rotated back and forth approximately $\pm 3^\circ$. Under this load the wheel support flexure struts (see Fig 2 of Ref. 1) were slightly bowed as measured by a straight edge. Upon removal of the ballast weights, the bowing disappeared. It was observed that

under a small load of about 100 N there was a slight asymmetry of the wheel support housing with respect to the flexure struts, which implied that the wheel had rotated a small amount in order to contact the track perfectly. This means that the strut was slightly bowed initially and that additional load increased the amount of bowing. The calculated axial load in each of the four flexure struts was 5642 N and the calculated critical load was 7592 N. Thus each strut was loaded to $5642/7592 = 74\%$ of its critical buckling load without any visible damage. Notice that this load of 19547 N on the wheel was $19,547/5560 = 3.51$ times the nominal load caused by antenna deadweight.

An examination of the wheel rolling surface and the wear strip disclosed no visible damage, nor was there any apparent damage to the portland cement grout beneath the track in the vicinity of the overloaded wheel.

V. Evaluation of the Fretting

There are many conflicting effects in the phenomenon of fretting corrosion, therefore it is not possible to know whether it would tend to be more or less severe on the prototype than on the model. The following conditions would be different

- (1) The time required to accumulate 67471 turns of the prototype antenna would be approximately 67 years, whereas this number of model turns was made in 180 hours
- (2) The amplitude of relative motion between the bottom surface of the wear strip and its support ring would be more on the prototype, namely, it would be $1/\lambda$ times the model amount
- (3) The model tests were conducted in a temperature- and humidity-controlled room, whereas the prototype would be exposed to the weather
- (4) The surface finishes of the prototype may be different from those of the model.

Since every lubricant used at the model wear strip interface was substantially effective in reducing the amount of fretting, and some virtually eliminated fretting, it is believed that some of these would be effective on the prototype. It is recommended that metallic platings not be used because their effectiveness is sacrificial in nature and would eventually disappear.

In order to retain lubricative substances at the interface, the edges of the wear strips, including the mitered ends, should be sealed to the support ring with a flexible material such as silicone rubber. Such a seal would also exclude foreign abrasives and moisture.

VI. Methods of Retaining the Support Ring

The method of retaining the model support ring, or track, is described in Ref. 1. Circumferential walking was prevented by four equally spaced tangential links. One end of a link was bolted to the outer edge of the support ring and the other end bolted to the concrete foundation. It was intended that the grout placed against each side of the support ring cross section (see Fig. 6) would resist side forces induced by slightly misaligned wheels. The fact is that a section of the outer grout shoulder, approximately 80° in angular span, was broken away and allowed the track to be displaced outward by as much as 1.4 mm. It is now believed that a portland cement shoulder is not suitable for retaining the support ring against these induced side forces nor against radial forces caused by a temperature difference between the support ring and concrete foundation. The latter force was not present on the model but would exist on the prototype.

It is believed that any of the three following methods would successfully retain the support ring of the prototype. The first method employs 16 equally spaced tangential links which connect the support ring to the concrete foundation. These links prevent circumferential walking, allow free radial expansion from temperature differences, and restrain the support ring against side forces induced by any combination of misaligned wheels. A structural analysis of this link arrangement has been made and will be reported separately. The results show that all stresses and displacements will be small. This method does permit small relative radial displacements between the support ring and the grout beneath it, which may be deemed unacceptable.

In order to eliminate the relative movement between the support ring and grout, the second method (see Fig. 7) has a thin sole plate between the support ring and the grout. The bottom surface of the sole plate must be keyed to the grout with both circumferential and radial keys. Sixteen tangential links would be used just as in the first method. Then the relative radial movement would be between the support ring and sole plate. The use of a sole plate requires that the bottom of the support ring be flat and that the sole plate be pressed against it firmly before the grout is packed into place. Another disadvantage of this method and of the first method is that the tangential links must be attached to steel plates embedded in the concrete foundation and not attached directly to the concrete because the bearing stresses would be too great. Also there is the possibility of fretting corrosion occurring between the sole plate and support ring.

The third method (see Fig. 8) would eliminate all the tangential links and employ circumferential and radial keys welded to the bottom of the support ring. The circumferential

key would be at the bottom center of the support ring. It would serve to prevent radial movement between the support ring and grout whenever there is a temperature difference between the two. Sixteen equally spaced radial keys would prevent track walking. A stress analysis shows that the additional grout stresses caused by the key forces could be made sufficiently small with keys of moderate size.

VII. Wear Strip Mitered Joints

The top surface of the support ring was originally smooth and flat and the thicknesses of adjacent wear strips were the same to within 0.005 mm. When the model first rotated at the rate of 375.55 revolutions per hour, the noise it produced was very low and was described as a gentle hum. However, the noise level increased rapidly with the elapse of testing time. At the end of the tests the noise was described as a loud drumming sound. When the model was slowly pushed by hand it was observed that the noise level was fairly low except when a wheel was traversing a mitered joint, at which time a clanking sound was made and the wheel assembly could be seen to wobble.

Early examination of the mitered joints showed that a peening action was occurring. As the effects of the peening increased, the noise and wheel wobble increased. It is now believed that there is an effective step at the joint, even though the adjacent wear strips are of the same thickness, caused by elastic deformation. Whether peening occurs probably depends upon whether yield stresses are exceeded. It will be shown in the following discussion that the material stresses during the traversing of a joint are dependent upon the velocity and that the relatively high speed model intensified these stresses. It is expected that the amount of wear strip peening near the joints would be much less or nonexistent on the prototype.

In order to obtain the velocity scaling factor needed, a simpler joint is analyzed, namely, one which is perpendicular to the track. In Fig. 9 the wheel is shown at the beginning and end (solid and dashed lines respectively) of a joint traverse. Also indicated in Fig. 9 is the circular arc path of the wheel center and its rectangular coordinates s and y . For very small values of the ratio of step height h to wheel radius R , the wheel center path can be approximated closely by the sinusoid

$$y = h \sin \frac{\pi}{2\sqrt{2Rh - h^2}} S \quad (7)$$

Let

$$S = VT \quad (8)$$

where V is the constant horizontal velocity of the wheel center and t is the time. The substitution of (8) into (7) gives the vertical position coordinate of the wheel y in terms of the time t , namely,

$$y = h \sin \frac{\pi V}{2\sqrt{2Rh - h^2}} t = h \sin \Omega t \quad (9)$$

Figure 10 shows the single-degree-of-freedom model to be analyzed. In this model the step is assumed to be perpendicular to the motion of the wheel, whereas on the physical model the step was formed by a 45° mitered joint. The main reason for analyzing the simpler condition is to obtain proper velocity scaling factors, which may be used to indicate that the prototype will have lower stresses during a joint traverse, even though the static stresses in the prototype and model are the same.

The equation of vertical motion of the dynamic model shown in Fig. 10 is

$$M\ddot{x} = -K(x - h \sin \Omega t) \quad (10)$$

where \ddot{x} is the acceleration of mass M

Defining the natural frequency ω as

$$\omega^2 = \frac{K}{M}, \text{ Eq. (10) becomes} \quad (11)$$

$$\ddot{x} + \omega^2 x = \omega^2 h \sin \Omega t \quad (12)$$

When the initial conditions are zero, the solution of (12) is

$$\ddot{x} = \frac{h\omega}{\omega^2 - \Omega^2} (-\Omega \sin \omega t + \omega \sin \Omega t) \quad (13)$$

The vertical velocity and accelerations of mass M are as follows

$$\ddot{x} = \frac{h\omega^2 \Omega}{\omega^2 - \Omega^2} (-\cos \omega t + \cos \Omega t) \quad (14)$$

$$\ddot{x} = \frac{h\omega^2 \Omega}{\omega^2 - \Omega^2} (\omega \sin \omega t - \Omega \sin \Omega t) \quad (15)$$

The valid range of time t for Eqs (13) (14) and (15) is

$$0 \leq t \leq \frac{\sqrt{2Rh - h^2}}{V} \quad (16)$$

For the application of concern, namely, azimuth bearing wheels for large antennas, the wheel mass is very small in comparison to the antenna mass associated with one wheel M , so that the assumption of a massless wheel is justified. The physical model which was tested had large ballast masses almost directly above a wheel. The flexure plates supporting the wheel (see Fig. 2 of Ref. 1) constitute most of the flexibility in the system, therefore this single-degree-of-freedom mathematical model is suitable for obtaining approximations for the real accelerations. The force between the wheel and track during a joint traverse may be obtained by multiplying \ddot{x} by the mass M since the wheel is assumed to be massless.

The wheel velocity of the prototype V_p , which would produce the same stresses during a joint traverse as was produced in the physical model, is obtained in the following discussion which is based upon the assumptions that the model and prototype static stresses are identical and that the joints are geometrically similar. The first assumption requires that the prototype mass M_p be equal to the model mass M_m divided by λ^2 , where λ is the ratio between corresponding model lengths and prototype lengths. The subscripts m and p refer to model and prototype respectively.

The foregoing assumptions are

$$M_p = \frac{M_m}{\lambda^2} \quad (17)$$

$$h_p = \frac{h_m}{\lambda} \quad (18)$$

$$R_p = \frac{R_m}{\lambda} \quad (19)$$

As a test will show, the required prototype velocity is

$$V_p = \frac{V_m}{\sqrt{\lambda}} \quad (20)$$

Equation (15) refers either to model or prototype. By writing \ddot{x}_p in terms of model parameters, all scale factors will cancel if the prototype velocity, as given by Eq. (20), is correct.

The spring constant K was obtained by determining the deflection of the wheel support flexures as shown in Fig. 2 of Ref. 1. The result is

$$K = \frac{3AE}{l} \quad (21)$$

where A is the cross sectional area of one flexure element, E is the elastic modulus, and l is the length of each flexure element

Assuming that the E 's are the same for model and prototype and that the two are geometrically similar, K_p may be expressed

$$K_p = \frac{3 \frac{A_m}{\lambda^2} E_m}{\frac{l_m}{\lambda}} = \frac{3 A_m E_m}{l_m} \frac{1}{\lambda} = \frac{K_m}{\lambda} \quad (22)$$

Substituting (22) and (17) into (11) there is obtained

$$\omega_p^2 = \frac{K_p}{M_p} = \frac{K_m \lambda^2}{\lambda M_m} = \omega_m^2 \lambda \quad (23)$$

From Eq (9) the expression for Ω^2 is

$$\Omega^2 = \frac{\pi^2 V^2}{4(2Rh - h^2)} \quad (24)$$

Substitute (18) (19) and (20) into (24) and obtain

$$\Omega^2 = \frac{\pi^2 V_m^2}{\lambda^4 \left(\frac{2R_m h_m - h_m^2}{\lambda^2} \right)} = \Omega_m^2 \lambda \quad (25)$$

Now substitute (18) and (23) into equation (15) and obtain

$$\ddot{x}_p = \frac{\frac{h_m}{\lambda} \omega_m^2 \lambda \Omega_m \sqrt{\lambda}}{\omega^2 \lambda - \Omega_m^2 \lambda} [\omega_m \sqrt{\lambda} \sin \omega_m \sqrt{\lambda} t_p - \Omega_m \sqrt{\lambda} \sin \Omega_m \sqrt{\lambda} t_p] \quad (26)$$

The substitution of (18) (19) and (20) into (16) yields

$$t_p = \sqrt{\frac{2R_m h_m - h_m^2}{\lambda^2}} \frac{\sqrt{\lambda}}{V_m} = \frac{t_m}{\sqrt{\lambda}} \quad (27)$$

The substitution of (27) into (26) yields

$$\ddot{x}_p = \frac{h_m \omega_m^2 \Omega_m}{\omega_m^2 - \Omega_m^2} [\omega_m \sin \omega_m t_m - \Omega_m \sin \Omega_m t_m] \quad (28)$$

Thus the \ddot{x}_p value, or acceleration, of the prototype as given by (28) is identically equal to (15) with subscripts m . Since (15) applies either to model or prototype and since (28) is equal to (15), it is clear that when the prototype velocity is as given by Eq (20), the prototype acceleration is equal to that of the model

The acceleration as given by (15) will now be evaluated for the model test conditions. The following values of the parameters are

$$h = 0.051 \text{ mm}$$

$$R = 25 \text{ mm}$$

$$V = 1.117 \text{ m/sec.}$$

$$K = \frac{3AE}{l} = \frac{3(0.0000242)(20)10^{10}}{0.057}$$

$$= 254,631,579 \text{ N/m}$$

$$M = \frac{5560}{9.8} = 567 \text{ kilograms}$$

$$\omega^2 = \frac{K}{M} = 449,085 \text{ rad/sec}^2$$

$$\omega = 670 \text{ rad/sec}$$

$$\Omega^2 = \frac{\pi^2 V^2}{4(2Rh - h^2)} = \frac{\pi^2 (1.117)^2}{4[2(0.025)(0.000051) - (0.000051)^2]}$$

$$\Omega^2 = 1,208,506 \text{ rad/sec}^2$$

$$\Omega = 1099 \text{ rad/sec}$$

$$t_{max} = \frac{\sqrt{2Rh - h^2}}{V} = \frac{\sqrt{2(0.025)(0.000051) - (0.000051)^2}}{1.117}$$

$$t_{max} = 0.00142 \text{ sec}$$

The substitution of the above parameters into Eq. (15) yields

$$\ddot{x} = \frac{0.000051(449085)1099}{449085 - 1208506} [670 \sin 670 t - 1099 \sin 1099 t] \quad (29)$$

$$\ddot{x} = -0.033 [670 \sin 670 t - 1099 \sin 1099 t] \quad (30)$$

The maximum absolute value of the bracket term of (30), within the time range zero to 0.00142 second, is approximately 581 rad/sec. Therefore the maximum value of \ddot{x} is

$$\ddot{x} = -0.033 [-581] = 18.99 \text{ m/sec}^2 \quad (31)$$

The additional wheel to track force, F_D , caused by this acceleration is obtained by multiplying it by the mass M , obtaining

$$F_D = M\ddot{x} = 567(18.99) = 10767 \text{ N} \quad (32)$$

The total wheel force is obtained by adding the static force F_s to the dynamic force F_D , obtaining

$$F_{total} = F_s + F_D = 5560 + 10767 = 16327 \text{ N} \quad (33)$$

The ratio of F_{total} to F_s is

$$\frac{F_{total}}{F_s} = \frac{16327}{5560} = 2.93 \quad (34)$$

The wheel velocity of a certain prototype antenna is 0.05 m/sec. If the model is considered to be a 1/12-scale model of this prototype, the model wheel velocity would be, from Eq. (20)

$$V_m = \sqrt{\lambda} V_p = \sqrt{\frac{1}{12}} 0.050 = 0.0144 \text{ m/sec} \quad (35)$$

The appropriate value of Ω^2 , is, from Eq. (24),

$$\Omega_1^2 = 202 \quad (36)$$

$$\Omega_1 = 14.2 \quad (37)$$

The appropriate value of the time range is from Eq. (16)

$$t_{1_{max}} = 0.111 \text{ sec} \quad (38)$$

The value of the acceleration \ddot{x}_1 , is obtained from (15)

$$\ddot{x}_1 = \frac{0.000051(449085)14.2}{449085 - 202} [670 \sin 670 t - 14.2 \sin 14.2 t] \quad (39)$$

$$\ddot{x}_{1_{max}} = 0.000724 [665] = 0.482 \text{ m/sec}^2 \quad (40)$$

The additional wheel to track force, F_{D1} , caused by this acceleration is

$$F_{D1} = M\ddot{x} = 567(0.482) = 273 \text{ N} \quad (41)$$

The total wheel force is obtained by adding the static force F_s to the dynamic force, F_{D1} , obtaining

$$F_{1_{total}} = F_s + F_{D1} = 5560 + 273 = 5833$$

The ratio of $F_{1_{total}}$ to F_s is

$$\frac{F_{1_{total}}}{F_s} = \frac{5833}{5560} = 1.049 \quad (42)$$

This value, which pertains to a model run at the appropriate speed, is to be compared to the equation (34) value of 2.936, which pertains to the actual model running speed. At the appropriate speed the static loading is increased by 4.9%, whereas at the speed at which the model was run, the static loading was increased by 193%.

The foregoing ratios of total load to static load were derived for a zero gap step perpendicular to the track. It is reasonable to believe that the same force ratio would apply to the case of a step in a mitered joint, although the stresses of the mitered joint are probably quite different. This may be seen by referring to Fig. 11. In Fig. 11(a), which shows the perpendicular step, the entire length of the wheel contact area will hit the step at one instant and the impact forces will be uniformly distributed over the length. In Fig. 11(b) is shown the mitered step. Here, one end of the contact area will first strike the step so that the impact forces will be concentrated near the edge of the wear strip. Peened areas were formed where the figure indicates. Figure 12 is a photograph which shows the worst peened joint. Most probably the yield stresses were exceeded in these areas. Generally the peened areas existed in pairs, which implies that a portion of the step was formed by elastic deflection of the loaded side of the joint. The major portion of the step, however, probably came from different thicknesses of adjacent wear strips.

In practice it is necessary to employ mitered joints. If finite gap perpendicular joints were used, the stresses at the free edges would be excessive and cause rapid wear and roughness.

The model was run at high rotational speeds in order to obtain a large number of turns in a few weeks of testing time. All static stresses matched those of the prototype. The wheel loads induced by steps at the wear strip mitered joints were intensified by the high testing speed. It is important to realize that had the model been run at the appropriate scaled speed, as given by Eq. (20), the wheel loads at the joints would have been much less, and it is believed these loads would have been approximately 4.9% more than for the loads existing at the joints at zero speed. The stresses near the joint at zero speed condition are not known, but are sure to vary considerably with the step size. Therefore it is important to maintain as small a step condition as is practical.

For the case of a zero step, the maximum shear stress τ can be estimated from Eq. (1) of Ref. 1, namely,

$$\tau = 0.179 \sqrt{\frac{F_s E}{DL}} \quad (43)$$

where D is the wheel diameter, L is its length, F_s is the static load from one wheel, and E is the common elastic modulus of the wheel and track.

For the case of low wheel velocities, where the effect of dynamic loading at a mitered joint is minor, it will be instructive to form the ratio of the calculated shear stress to the shear yield and then compare this ratio to corresponding ratios of known successful designs. The wear strip material of the model had a tensile yield stress of $105000 (10^4) \text{ N/m}^2$. The shear yield may be taken as a certain percentage of the tensile yield, say 50% of it, thus obtaining $52500 (10^4) \text{ N/m}^2$ as the shear yield of the material. Using this together with the calculated stress from Eq. (43), the following ratio is obtained, which is applicable for the model when running at the appropriate low speed.

$$\left[\frac{\tau}{\tau_{\text{yield}}} \right]_{\text{model}} = \frac{0.179}{52500 \cdot 10^4} \sqrt{\frac{5560 (20) 10^{10}}{(0.050) (0.0105)}} = 0.496 \quad (44)$$

The successful design with which the model will be compared is the radial bearing of the 64-m-diameter antenna which has been in service at DSS 14 for 15 years. Its wear strips have a material tensile yield of $69000 (10^4) \text{ N/m}^2$, or a calculated shear yield of $34500 (10^4) \text{ N/m}^2$. The wheel diameter, effective wheel length, and wheel force of this design are respectively 0.914 m, 0.101 m and 445000 N. Using Eq. (43) the following ratio is obtained.

$$\left[\frac{\tau}{\tau_{\text{yield}}} \right]_{\text{radial bearing}} = \frac{0.179}{34500 (10^4)} \sqrt{\frac{445000 (20) 10^{10}}{(0.914) (0.101)}} = 0.509 \quad (45)$$

Since there has been no damage to the edges of the wear strip mitered joints of the radial bearing assembly, and since its stress ratio as given by (45) is approximately the same as that of the model when the model runs at the appropriate low speed, it is concluded that had the model completed its large number of turns at low speed, there would not have been any peening at the ends of the wear strips. It should be emphasized that even after all the high speed turns the noisy model was operable in spite of the peening at the joint ends.

VIII. Conclusions

From the foregoing evaluation of the model tests the following conclusions are made:

- (1) There is a tendency for a large-diameter circular track to move circumferentially and radially. If portland-cement-type grout is used for supporting the track, it is necessary to use auxiliary restraint devices to prevent the movement. It is believed that any of the three methods described would be satisfactory. Auxiliary restraint devices may not be required if another grout material, such as an epoxy-bonded grout, is used, but since this is not known, it is recommended that the auxiliary restraint devices be used with any grout material.
- (2) The hardnesses of the wheels and wear strips apparently are sufficient to insure a long service life when used in conjunction with the wheel loading and wheel suspension system described. A detailed study of the test conditions summarized in Table 3 is necessary in order to make statistical predictions about the expected life of a prototype. The third series of tests, which was run at increased wheel loadings to compensate for scale effect, survived 14361 turns. It may be concluded that a three-wheel prototype would survive this number of turns.
- (3) The wheel suspension survived an overload of 3.51 times its nominal load.

- (4) The torque required to overcome the rolling resistance of the wheel can be estimated sufficiently well by using Eq. (8) of Reference 1.
- (5) The traction capacity was found to be consistent with that from other sources
- (6) The infrequent transverse sliding of a nonrotating wheel across the track produces only negligible damage when the contact stresses and material properties are as herein stipulated.
- (7) Fretting corrosion occurred on the model at the interface between the wear strips and the support ring. Two lubricants were found which almost eliminated fretting. Since it is probable that fretting would occur on a nonlubricated prototype, it is recommended that suitable lubricants be provided.
- (8) The peening which occurred near the model wear strip ends is not expected to occur on the prototype because of velocity scale effect.

Reference

- 1 McGinness, H, "Antenna Azimuth Bearing Model Experiment," *The DSN Progress Report 42-53*, Jet Propulsion Laboratory, Pasadena, Calif, October 15, 1979

Table 1. Qualitative comparison of lubricants used in controlling fretting corrosion at interface (the greatest amount of fretting from this group was small in comparison to the condition of no lubricant)

Wear strip number	Lubricant	Condition of interface after 49 13 hours of second series test,	Relative ranking class
2	Cortec	Very slight fretting, least amount	1
14	Silver Goop	Very slight fretting, least amount	1
10	Chassis grease	Small amount of fretting	3
3	Wheel brg. grease	Small amount of fretting	3
7	SAE 30 oil	Moderate amount of fretting	4
1	Darina No 2	Moderate amount of fretting	4
16	Mo S2	Considerable, but less than graphite	4
4	Graphite	Considerable fretting, greatest amount	5
5	Graphite	Considerable fretting	5
6	Graphite	Considerable fretting	5

Table 2 Qualitative comparisons of lubricants used in controlling fretting at interface

Wear strip number	Lubricant	Condition of interface after an additional 32 41 hours of second series test	Relative ranking class
2	Silver Goop ^a	Very slight amount of fretting, best class	1
14	Silver Goop	Very slight amount of fretting, best class	1
20	Silver Goop	Very slight amount of fretting, best class	1
10	Silver Goop ^a	Slight amount of fretting, next best class	2
17	Chassis grease	Considerable fretting, worst class	5
3	Silver Goop ^a	Small amount of fretting	3
13	Wheel brg grease	Considerable fretting	5
7	Silver Goop ^a	Small amount of fretting	3
18	SAE 30 oil	Considerable fretting	5
11	SAE 30 oil	Considerable fretting	5
19	SAE 30 oil	Considerable fretting	5
7-2/3	SAE 30 oil	Considerable fretting	5
1	Silver Goop ^a	Moderate fretting	4
9	Darina no 2	Considerable fretting	5
16	Silver Goop ^a	Moderate fretting	4
15	Dry MoS2	Considerable fretting	5
4	Silver Goop ^a	Small amount of fretting	3
5	Silver Goop ^a	Small amount of fretting	3
6	Silver Goop ^a	Small amount of fretting	3

^aThese employed a different lubricant during preceding 49 13 hours of testing, see table 1

Table 3. Summary of test conditions

	Load per wheel, N	No. of wheels	Lubrication between wheel and wear strip	Wheel misalignment, arcmin	Total running time, hr	Turns of antenna model
First series	5560	6	Products of fretting corrosion	Less than 0.50	40.60	15247
Second series	5560	6	Inadvertent lub. from substances beneath wear strip. Also polyoil	Less than 0.50	81.54	30622
<u>Wheels replaced with new ones</u>						
Third series	8913 8749	3	Polyoil plastic wiper	Less than 0.50	38.24	14361
Fourth series	5674	3	Polyoil plastic wiper	40	19.28	7241

Total turns of antenna model with first set of wheels was 45870. Each wheel of first set made 3256770 revolutions about wheel axis.

Total turns of antenna model with second set of wheels was 21602. Each wheel of second set made 1533742 revolutions about wheel axis.

Product of antenna model turns and number of wheels was 340026.

No evidence of fatigue damage to wheels.

Only wear strip No. 15 sustained spalling damage, and this occurred during fourth series of tests.

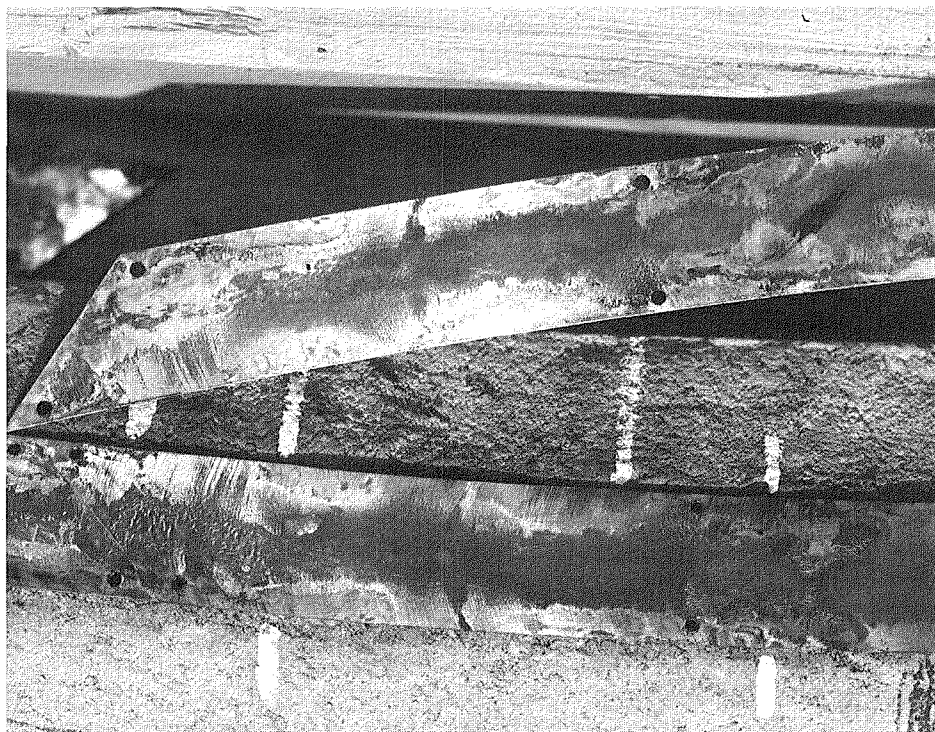


Fig. 1(a). Removed wear strip beside its support ring, showing fretting corrosion after 40.6 hours of running

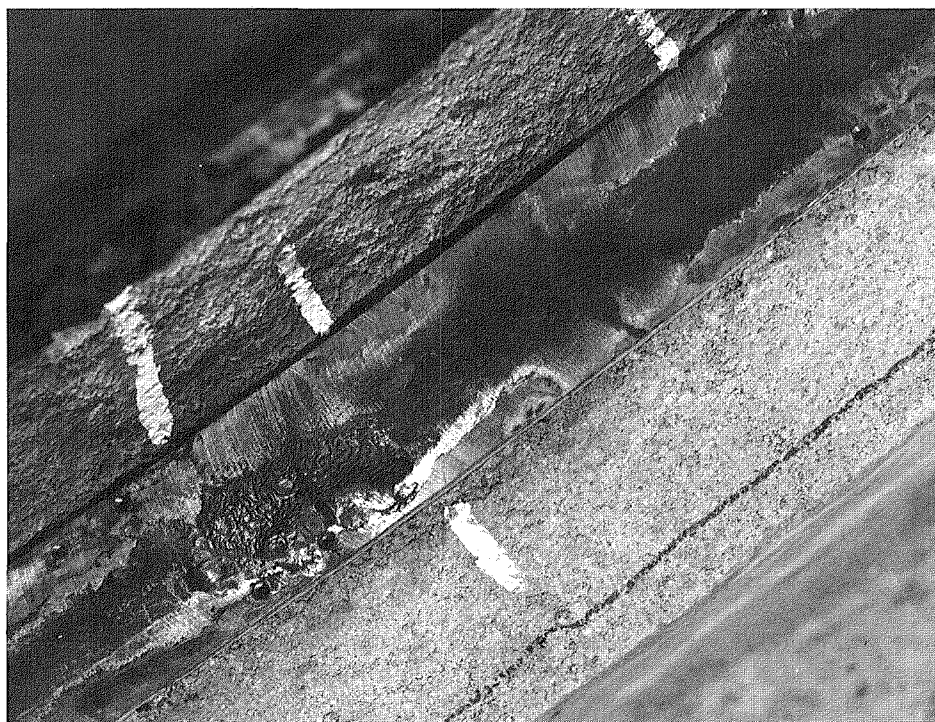


Fig. 1(b). Enlarged view of portion of Fig. 1(a)

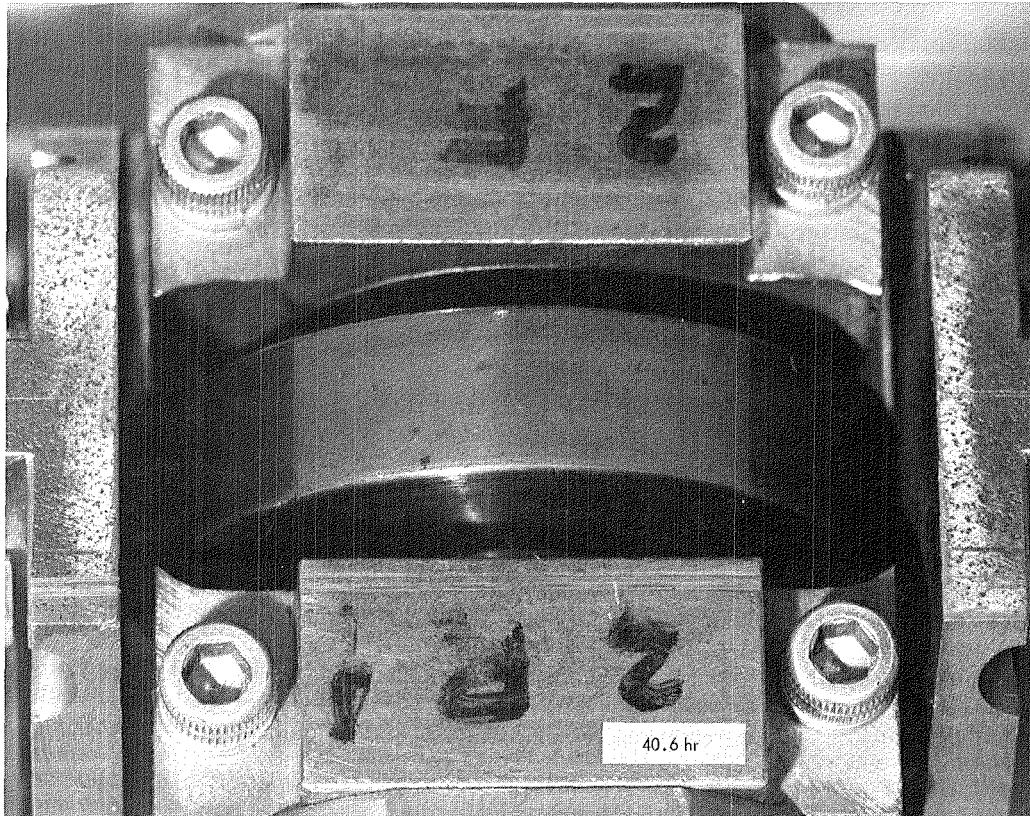
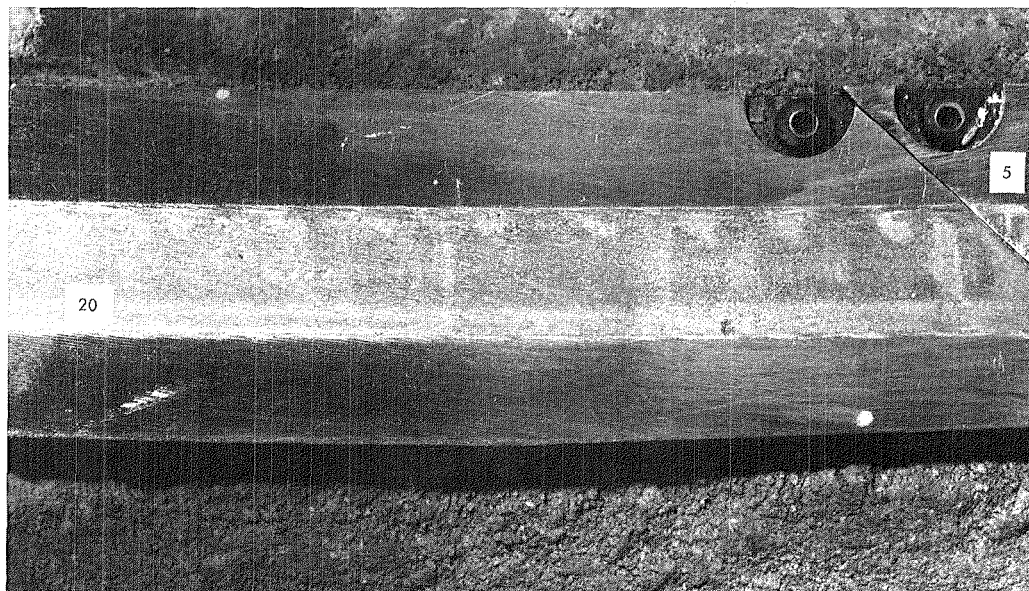


Fig. 2. Wheel surface after 40.6 hours of running



C AFTER 40.6 hr

Fig. 3. Wear strip top surface after 40.6 hours of running

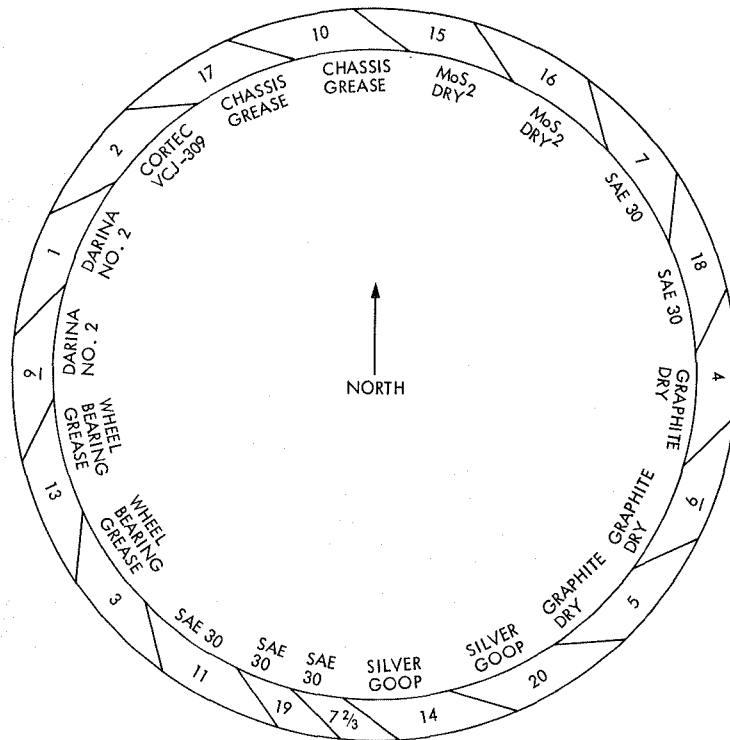


Fig. 4. Wear strip identification numbers and disposition of lubricants between wear strips and support ring

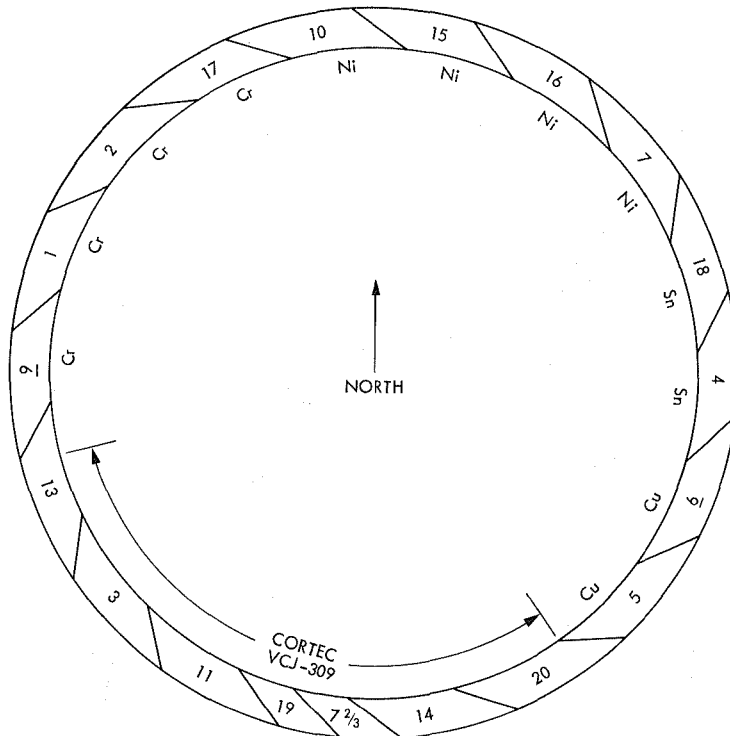


Fig. 5. Wear strip identification numbers and lubricant disposition for third series of tests

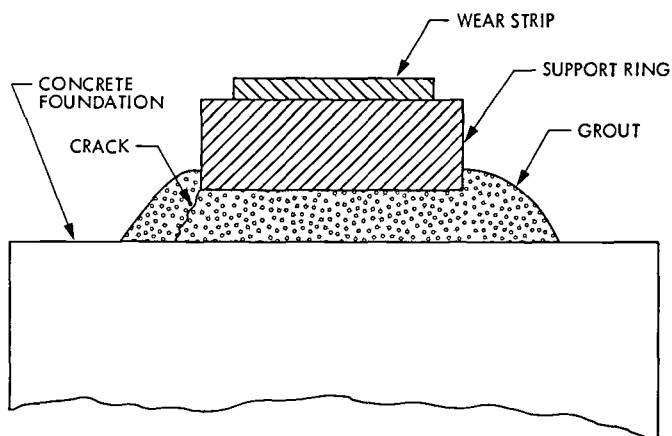


Fig. 6. Cross section of track showing cracked grout shoulder on outer side

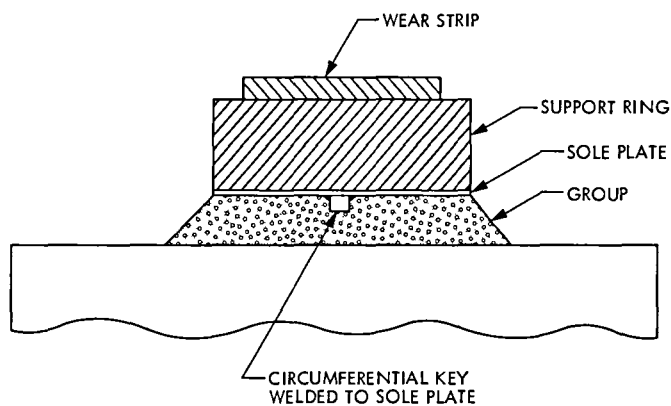


Fig. 7. Cross section of track showing addition of sole plate keyed to grout, method 2

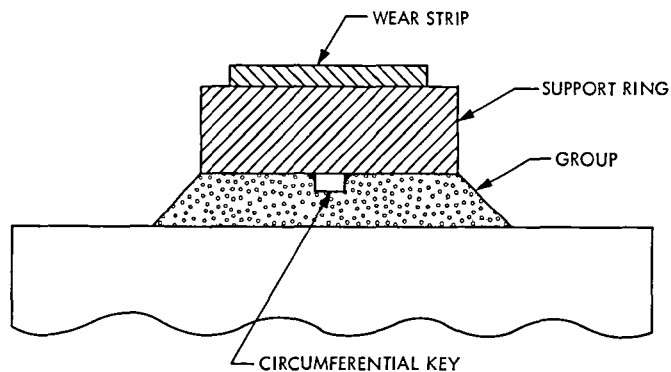


Fig 8 Cross section of track showing circumferential key, method 3

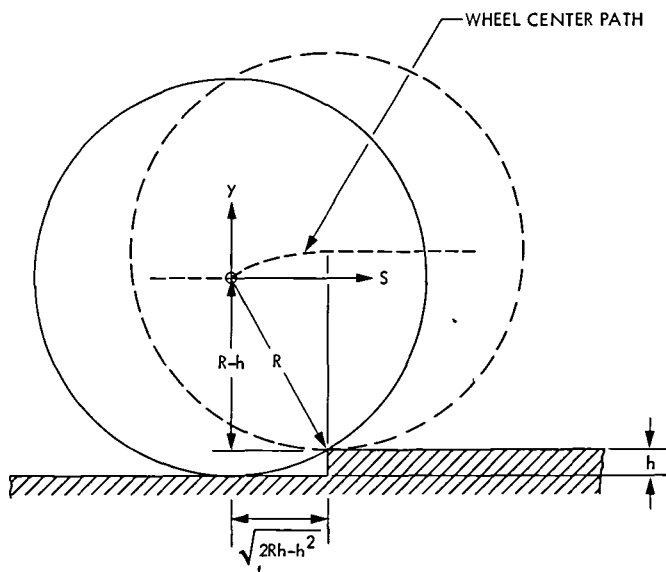


Fig 9 Configuration of wheel against step at wear strip joint

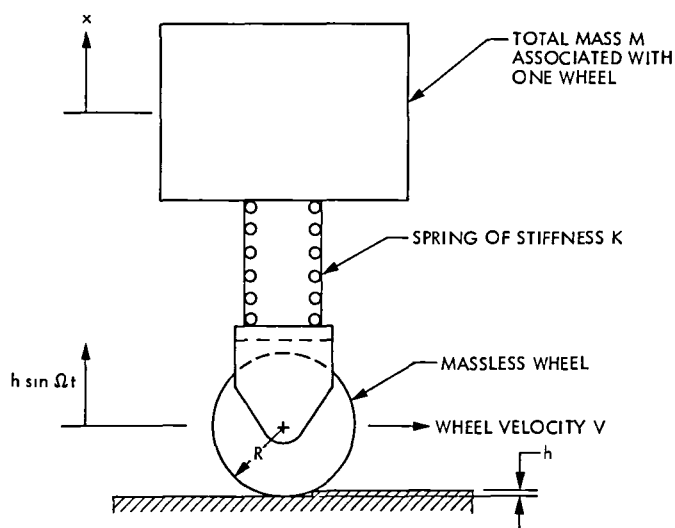


Fig 10. The single-degree-of-freedom model analyzed

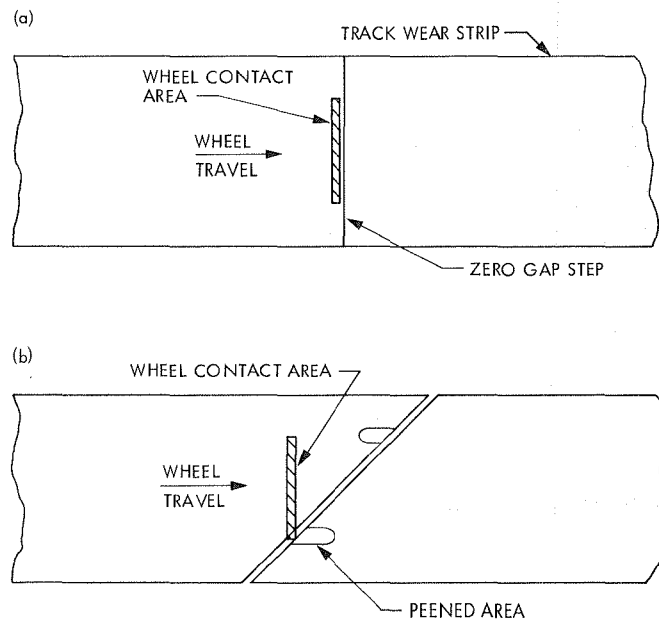


Fig. 11. Comparison of zero gap step perpendicular to track with a finite gap mitered step

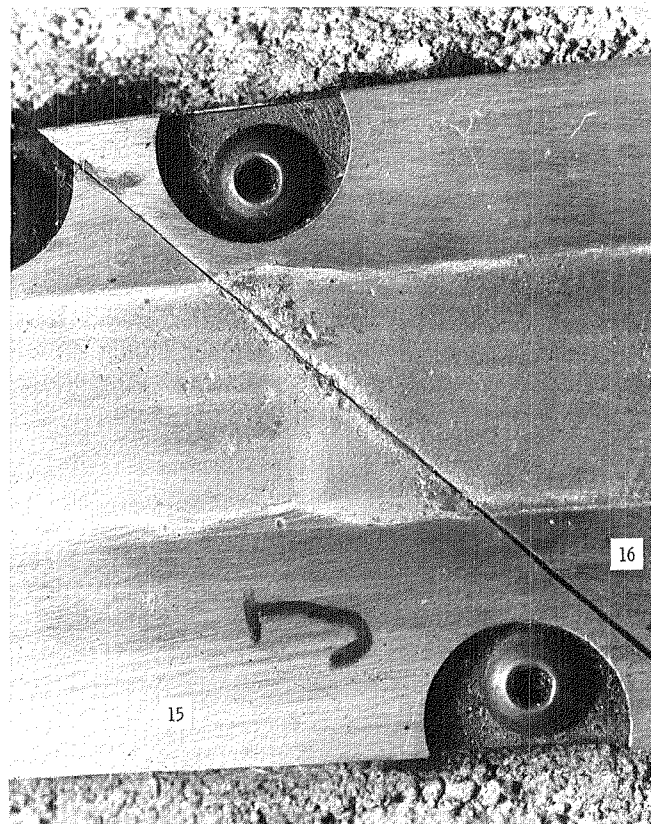


Fig. 12. Photograph showing peened areas at mitered point

Operational VLBI Clock Synchronization and Platform Parameter Determination

J V LuValle and R D Shaffer
Control Center Operations Section

M G Roth, T M Eubanks, and P S Callahan
Tracking Systems and Applications Section

The operation and results of the Block I VLBI system for clock synchronization and platform parameter determination are described. The history of system performance from July 1980 to August 1981 is given. The steps in producing clock sync and platform parameters are outlined. It is shown that the VLBI data provide useful monitoring of the Deep Space Network Frequency and Timing System performance and earth orientation.

I. Introduction

The Block I VLBI (Very Long Baseline Interferometry) system (Ref. 1) has been developed at JPL to provide near-real-time Deep Space Station (DSS) clock synchronization (clock sync), platform parameter determination (UT1 and polar motion — UTPM), and spacecraft navigation. From July 1980 to August 1981 system development was carried out concurrently with observations designed to measure clock sync and UTPM. This report describes the steps involved (see Fig. 1) in conducting and processing the observations. The history of system development and performance during this period is discussed. The disposition of 101 passes scheduled from July 14, 1980, to August 29, 1981, is given in Table 1. Clock sync results for December 1980 to August 1981 are given in Table 2 and Figs. 2 and 3.

II. Observation Scheduling

Beginning in 1980, weekly VLBI observing sessions were scheduled on both the California-Spain and California-Australia baselines for the purpose of making clock sync and UTPM measurements and for Block I (Ref. 1) system development and checkout. The 64-m DSS was used because Block I equipment is available there, and the recording bandwidth (250 kHz) of the Block I system requires the sensitivity of the large antennas. In most cases both baselines were observed in the same 24-hour period. Estimation of clock parameters requires only single baseline observations, but UTPM estimates require two baselines to separate all three parameters.

For each baseline, 7 to 13 extragalactic radio sources (EGRS) from the radio source catalog were observed for

200 seconds each. Source selection and pointing predicts were generated in special purpose software. A tape of pointing predicts was given to Operations for reformatting and transmission by high speed data line (HSDL) to the DSS. For most of the 1981 observations, the scheduling of sessions was governed mainly by DSS availability. However, during the period from DOY 136 through 200, a sidereal schedule was employed, which permitted the same sources to be observed each week for 10 consecutive weeks. The sidereal schedule consisted of a set of 12 observations which were done in the same sequence each time. The observing time shifted at the sidereal rate (approximately 4 min/day) so the same sources could be used. The sidereal schedule proved to be very helpful in the identification of systematic problems.

Standard DSS configuration and observing procedures were developed. These procedures were the Deep Space Network Operations Plan for the Block I VLBI System written by the Mission Coordination Group of the Control Center Operations Section and the Deep Space Station Standard Operating Procedure Very Long Baseline Interferometry (VLBI) System Block I written by the System Support Group of the Deep Space Network System Support Section. These procedures and the automated pointing predicts resulted in relatively smooth operations. Initially data were acquired in four S-band and four X-band bandwidth synthesis (BWS) channels with a maximum spanned bandwidth of 30 MHz in each band. To improve signal-to-noise ratios, the configuration was changed to three BWS channels per band with a maximum span of 40 MHz on 1981 DOY 107. The BWS technique provides the accuracy of a large spanned bandwidth (30-40 MHz) while recording only a small bandwidth (250 kHz). This is accomplished by sequentially sampling 250-kHz channels spaced across the desired spanned bandwidth. The channels are processed separately and the results combined to produce an effective bandwidth comparable to the spanned bandwidth. Occasional mistuning of the S-band travelling wave maser resulted in loss of one of the outer channels, but this did not greatly affect the overall quality of the results since X-band data were used for all solutions to date.

A total of 101 passes (Table 1) were scheduled from July 14, 1980, through August 29, 1981. Of these passes 72 were conducted, 41 produced clock sync results, and 32 produced useable data for UTPM solutions. Table 1 indicates at what point a pass failed from schedule to solution. A more important statistic is the success rate of passes after data processing was moved to the Block I VLBI Processing System (VPS) on March 28, 1981 (DOY 087). Since then, 44 passes were scheduled. Of these 11 were cancelled, 2 had data lost, and 24 produced cross correlation fringes (24 of 31). Of the 24 sessions which gave fringes, 22 produced clock sync results

and 17 were useable for UTPM solutions. This improvement in overall system performance is very encouraging.

During 1981, passes were scheduled by the Voyager Project to minimize conflicts. A total of 46 sessions were conducted through August 2, 1981, at which time the Voyager 2 Saturn encounter and the DSS 14 bearing problems terminated activities. Observing sessions will begin again after DSS 14 scheduled downtime in 1981 October.

III. Data Acquisition

Data are recorded on tape at each station by the Occultation Data Assembly (ODA) which is shared by the VLBI and Radio Science Subsystems. The recording rate is 500 kbits/sec. Currently, data are replayed over wide-band data lines (WBDL) at 56 kbits/sec directly onto the discs of the Network Operations Control Center (NOCC) VLBI Processing Subsystem (VPS) (see Fig. 1). Before 1981 DOY 086, data were obtained from intermediate data record (IDR) tapes, which resulted in some of the problems discussed below. Since the playback rate is 1/9 of the record rate, it takes approximately 6 hours to play back the 40 minutes of data from each pass.

In the process of implementing the first operational VLBI system, numerous problems in data acquisition and playback were solved in the past year. Two operational difficulties were resolved. First, during acquisition the NOCC had little visibility into DSS activities. Use of the Network Radio Science real-time monitor (RTM) partially solved this problem. Second, there were no personnel to support VLBI acquisitions and playbacks in real-time. NOCC controllers were trained and given detailed procedures (as discussed above) to cover these activities.

A number of problems in data acquisition were difficult to detect at first because of the very long turnaround time caused by doing all the processing at Caltech from IDRs. The slow detection of problems and lack of feedback to the stations resulted in additional bad data. Now, rapid turnaround with the VPS gives better DSS awareness and NOCC control of VLBI passes. Weak and poor fringe results were traced in part to antenna pointing errors. This was largely corrected by implementing an antenna pointing program developed by R. Livermore of Tidbinbilla Deep Space Communication Complex, Australia. This program allows for entering pointing offsets automatically during tracking to significantly reduce errors due to antenna distortion and axis pointing errors. Recently, slew restrictions at DSS 14 due to its azimuth bearing problem have resulted in occasional loss of one or two sources per pass and the cancellation of several clock sync passes during Voyager 2 Saturn encounter. A temporary

solution currently being tested is to transmit the VLBI base-band signal from DSS 13 to DSS 14 by the microwave link in order to use the ODA at DSS 14. Other hardware has also caused some difficulties. In the initial implementation of the DSS Block I system, ODA tapes consistently failed to start on time. This has been largely solved by operational workarounds and program modifications. There also have been a small number of VPS and peripheral failures which have all been fixed.

An overall real-time verification of the VLBI observation is desirable. An important step in this direction is a real-time capability to validate the performance of the microwave and VLBI hardware. This capability is being developed by monitoring the phase calibration tones (phase cal) at the ODA. Phase cal tones are injected into the front end of the maser amplifier and can be detected in real-time at the ODA for system validation or later when the data are correlated. Phase cal can then be used to correct the VLBI data for dispersive effects in the instrumentation.

UTPM solutions require data from a number of EGRS distributed about the sky. Several problems have been associated with the source schedules. First, the slews between sources did not adequately take into account antenna prelimits, cable wrap, or keyholes. This problem was handled by manually plotting each acquisition on a stereo plot for each station involved and providing warning notes to the controller. Another problem, still being worked, is not having enough sources in the source catalog to provide the desired number of observations on the DSS 14/DSS 63 baseline within scheduling constraints.

In the playback area there were two major problems, both of which now appear to be solved. The first involved the lack of monitoring of the playback data when the normal Ground Communications Facility (GCF) system was used. This difficulty was solved by direct playbacks to the VPS discs. The VPS software provides playback monitor capability and maintains a log of the data received and of the numbers and types of errors. This also permits immediate validation of the GCF lines from the DSS to the VPS. The GCF IDR capability is used as a backup. A temporary problem of a significant percentage of IDR tapes being unreadable on the VPS was also solved by direct playbacks to the VPS, eliminating the additional processing step of reading IDRs onto disc. The cause of the bad IDRs was apparently physically bad tapes. Before direct replays to the VPS began, the problem could sometimes be overcome by regenerating the data onto good tapes.

Operational experience has shown that the inability of the ODA to do selective recalls makes playbacks clumsy if any significant line outages occur. When outages occur, replay has

to begin at the beginning of the ODA tape being played. Each ODA tape takes about 90 minutes to play back. A future ODA implementation will provide disc instead of tape data recording at the DSS, and include a selective recall capability.

IV. Correlation

The activities involved in correlation, from an operations viewpoint, are data verification, correlation control record generation, fringe search, and postcorrelation record (PCR) generation. During correlation there is some indication of whether there are fringes and phase cal tones, and the fringe quality for the pass can be estimated. Detailed results (fringe and phase cal amplitudes) on a channel-by-channel basis are best obtained from postcorrelation analysis.

Data verification simply involves verifying that data from both stations are on the VPS discs, and that enough ancillary data are available to create a correlation control record.

The control record generation step includes running "ANCEDIT," which generates a correlation control record (CCR) automatically from the ancillary data and other stored information. This CCR was manually examined and edited as necessary. If peculiarities were noted in the CCR, information on the pass was reviewed and correlation would be attempted with the unedited results. For example, if it appeared that both DSSs used a nonstandard setup, the reported frequencies rather than the requested values were used. Usually the information that needed to be edited was source locations or source names.

After the CCR was examined, the "PROCESS" mode of the correlation program was loaded, and a run on the data was made looking for fringes. If no fringes were found, the CCR was checked again, and the time offsets between the DSSs were verified. If this was all correct, a single channel on a strong source was correlated. If no fringes were found again, a complete PCR was run and passed on to the TEMPO (Time and Earth Motion Precision Observations) Team for further processing.

If fringes are found, the "FSEARCH" mode of the program was run to get the approximate relative DSS clock offset. All of the offset was assigned to DSS 14 in the CCR (this makes postcorrelation processing easier), and a continuous PCR of the pass was generated. A backup copy of the PCR was written to tape for archival purposes. The TEMPO Team was then informed of the availability of the PCR.

A small number of passes result in no fringes or very few fringes with no obvious reason for the poor results. Improved

phase calibration monitoring with the next ODA software revision should help in identifying this problem. Possible reasons for no fringes include (1) weak sources, (2) hardware failure, (3) an undetected error in setup (most likely incorrect frequency settings), (4) something as yet unsuspected. Passes where there have been poor results for unknown reasons are being investigated. It was generally found that a pass is either very bad or very good, i.e., no fringes or 80-100% fringes.

V. Postcorrelation Processing

A. Software Development

Beginning with 1981 DOY 086 all postcorrelation processing was done on the Block I VAX (Digital Equipment Corp Virtual Address Extension, hereafter simply VAX) computer. The four major steps in postcorrelation processing and the software modules involved are: (1) fringe phase tracking (PHASOR), (2) phase calibration tone tracking (PCAL), (3) bandwidth synthesis ambiguity resolution (FITDELAY), and (4) clock and UTPM estimation (MASTERFIT). These modules were transferred to the Block I VAX from the Caltech IBM 3032. Extensive tests and comparisons with the modules still running on the IBM 3032 disclosed several software errors. These errors were corrected and the VAX and IBM versions brought into agreement before processing was shifted to the VAX.

A significant software development effort was carried out on the VAX to provide the interface between the VPS and PHASOR/PCAL. Several utility routines were created for transferring PHASOR/PCAL output from the VAX to the IBM for testing and for estimation as MASTERFIT was the last module transferred. The modules PHASOR and PCAL are also used in processing spacecraft VLBI data (Differenced Downlink Only Range, DOR).

B. Clock Synchronization Results

Of the 46 observing sessions conducted in 1981, 30 produced useful data for clock synchronization. As shown in Table 2 the clock offset uncertainties ranged from 35 to about 200 nsec. These data were processed without phase calibration for instrumental delays. Therefore, a bias is expected in the reported clock offset values. Comparisons with the offsets measured by a travelling clock indicate that the bias is smaller than 1 μ sec on both baselines (Figs 2 and 3). Data from DOR passes are also included in the figures since observation of only one EGRS is needed to establish the clock offset. All VLBI-derived clock offsets are in good agreement. Comparisons of the VLBI data with the FTS records of clock behavior by S. C. Ward, Frequency and Timing System (FTS) System Cognizant Operation Engineer, gave independent confirmation of a number of known clock

anomalies. Figures 2 and 3 show the VLBI data with rates fitted (see below) and discrepant points indicated. The main causes of disagreement were (1) a loss of synchronization between the VLBI subsystem and the FTS epoch which gives an error of $N \times 200$ nsec (N an integer), and (2) a 1- μ sec retardation of the DSS 14 epoch measured by the FTS standards laboratory. The cause of the erratic behavior labelled 3 in Fig. 3 is not yet known. The steep slope at the beginning of Fig. 3 is due to the use of a Cs standard at DSS 63 from November 11, 1980, to January 14, 1981.

The long-term clock rate calculated from the slope of the offset data was found to be $0.51 \pm 0.009 \times 10^{-12}$ s/s on the DSS 14/43 baseline and $0.05 \pm 0.02 \times 10^{-12}$ s/s on the DSS 14/63 baseline. These values agree with the LORAN determinations of clock rate and are about 10 times more accurate. The clock rates from individual sessions have uncertainties ranging from 0.13 to 1.5×10^{-12} s/s. The averages over the clock rates for each baseline agree very closely with the long-term rates.

C. UTPM Estimation

Of the 46 observing sessions conducted in 1981, 21 produced useful data for UTPM estimation. This success rate can be contrasted with the 1980 performance, when only 11 of 38 passes produced UTPM solutions.

After fringe fitting, delay observables were constructed by bandwidth synthesis, while the delay rate was derived from a single channel fringe frequency. Solutions using the delay and delay rate data are obtained with the program MASTERFIT. Due to the limited number of sources observed during each session, the solve-for parameters were restricted to earth orientation parameters plus a first-degree-polynomial clock model and a term allowing for systematic errors between the delay and delay rate data. The solutions rely upon a priori knowledge of the baseline vectors and source positions. The tropospheric and ionospheric delays were modeled by tables of average zenith delays and elevation angle mapping functions. Since dual-frequency (S and X bands) data were recorded, it was possible to estimate the ionospheric delay from the data, but solutions done using SX data had, in general, larger root mean square (rms) residuals than corresponding X-band solutions.

The original TEMPO solution strategy was, when possible, to combine data from paired observing sessions and to solve for all three UTPM parameters. If data from only one session were available, solutions were done for two of the three standard UTPM components. Such single baseline solutions rely upon outside information (the Bureau Internationale de l'Heure (BIH) Circular-D) for the unadjusted component

Errors or biases in the unadjusted component will map into the adjusted parameters, and the results may be corrupted. Since one purpose of these experiments was to provide results independent of the BIH, this was not a satisfactory method of solution. Even if data were available from both observing sessions, any changes in the earth orientation between the two sessions would corrupt the joint solution. In addition, any solution involving combined data obscures the contribution from the individual observing sessions and makes it hard to evaluate session performance.

Single baseline VLBI observations are insensitive to rotations about the baseline and thus cannot provide all three UTPM components without further constraints. Rotations of the earth about the baseline vector are equivalent to motions along a line parallel to the degenerate direction in the UTPM parameter space. If the UTPM parameter adjustments are constrained to lie in a plane normal to the degenerate direction, errors in the a priori value for the UTPM in the constrained direction will not affect the residuals or postfit parameter adjustments in the constraint plane. Solutions with such constraints are called normal plane solutions.

A program (Earth Rotation Analysis Of Variance (ERANOV)) was written to process normal plane solutions. ERANOV accepts as inputs the parameter adjustments, formal parameter errors, and correlation matrix resulting from a MASTERFIT normal plane solution and implements the solution analysis described above. The results obtained from the analysis of normal plane solutions consist of parameter adjustments describing a point in the constraint plane, together with the error ellipse associated with the point. A complete derivation of normal plane solutions will be published.

The overall results of carrying out the UTPM analysis on the data listed in Table 1 (X-band data only) are summarized below.

- (1) Sent to the BIH -- data from July 14, 1980, through March 1, 1981

15 observing sessions had usable results out of 53 scheduled and 39 conducted

154 delay and 153 delay rate observations in the 15 usable passes

RMS delay residual = 0.631 nanosec

RMS delay rate residual = 0.338 picosec/sec

- (2) In analysis -- data from March 9, 1981, through August 29, 1981

17 observing sessions had usable results out of 48 scheduled and 33 conducted

142 delay and 146 delay rate observations in the 17 usable passes

RMS delay residual = 0.985 nanosec¹

RMS delay rate residual = 0.748 picosec/sec¹

VI. Summary and Conclusions

The operation and results of the Block I Operational VLBI time synchronization and platform parameter determination system were outlined. Problems which arose during system startup were described. Most of the problems were overcome as operational equipment and software, in particular the VLBI Processing System on the Block I VAX, were implemented. Better operational procedures also contributed to improved data quality.

Useful clock sync and UTPM data were obtained for the period July 14, 1980, to August 2, 1981. The clock data were a factor 10 more accurate than other available data and clearly show known clock anomalies. UTPM data were sent to the BIH to aid in understanding the earth's orientation.

Thanks to the efforts of a large number of people at JPL, especially in the operations area and at the Deep Space Stations, the Block I VLBI system has developed into a powerful and well-functioning tool for the DSN to use in monitoring FTS performance and earth orientation.

¹These rms residuals should be regarded as preliminary. The cause of the increase in the rms residuals between the two data sets is being investigated.

Reference

- 1 Chaney, W. D., "The DSN VLBI System Mark IV-85," *TDA Progress Report 42-64*, pp. 61-76, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1981.

Table 1. Status of JPL Tempo Block I VLBI experiments

Explanation of the steps							
		Pass	= observations conducted				
		Corr	= correlated successful				
		PHASOR	= fringe fitting				
		FITDELAY	= ambiguity resolution				
		MASTERFIT	= least squares solution to data				
MASTERFIT solutions require at least 6 data points for good results Experiments without at least 6 X-band delay observations were cancelled Experiments with more S-band data may be used in the future							
X = step successfully completed							
C = processing cancelled at this step							
Baseline ID							
A = 14/43 experiment							
B = 14/63 experiment							
Date	ID/DOY	Pass	Corr	PHASOR	FITDELAY	MASTERFIT	Notes
07/14/80	B 196	X	X	X	X	X	Data sent to BIH
07/15/80	A 197	X	X	X	X	X	Data sent to BIH
07/19/80	B 201	X	X	X	X	X	Data sent to BIH
07/21/80	A 203	X	C				No fringes
07/25/80	B 207	X	X	X	X	X	Data sent to BIH
07/31/80	A 213	X	X	X	X	C	Not enough data
08/02/80	B 215	X	C				No fringes
08/05/80	A 218	X	X	C			No convergence
08/24/80	A 237	X	X	X	X	X	Data sent to BIH
08/24/80	B 237	X	X	X	X	X	Data sent to BIH
08/11/80	A 224	X	C				No fringes
08/11/80	B 224	X	C				No fringes
08/18/80	A 231	X	C				No fringes
08/30/80	A 243	C					Station problems
09/15/80	B 259	X	X	X	X	X	Data sent to BIH
09/23/80	A 267	X	X	X	X	X	Data sent to BIH
09/24/80	B 268	X	X	X	X	X	Data sent to BIH
09/30/80	A 274	X	X	X	X	X	Data sent to BIH
10/01/80	B 275	C					Pass cancelled
10/09/80	A 283	X	C				No fringes
10/10/80	B 284	X	X	X	X	C	Bad IDR data
10/16/80	A 290	X	X	X	X	X	Data sent to BIH
10/17/80	B 291	X	X	X	X	X	Data sent to BIH
10/23/80	A 297	X	C				Station problems
10/24/80	B 298	X	C				Playback problems
10/31/80	A 305	C					Pass cancelled
10/31/80	B 305	C					Pass cancelled
11/06/80	A 311	X	C				Playback cancelled
11/06/80	B 311	X	C				Playback cancelled
11/22/80	B 326	X	C				No fringes
11/22/80	A 327	X	X	X	X	X	Data sent to BIH
12/01/80	B 336	X	C				Unreadable IDRS
12/02/80	A 337	C					Pass cancelled
12/07/80	B 342	X	X	X	X	C	Not enough data
12/08/80	A 343	X	X	X	X	X	Data sent to BIH
12/13/80	B 348	X	X	X	X	X	Data sent to BIH
12/14/80	A 349	X	X	X	X	X	Data sent to BIH

Table 1 (contd)

Date	ID/DOY	Pass	Corr	PHASOR	FITDELAY	MASTERFIT	Notes
12/23/80	A 358	X	X	X	X	X	Data sent to BIH
12/23/80	B 358	C					Station problems
12/30/80	A 365	X	X	X	X		Ready to MFIT
12/31/80	B 366	X	X	X	X	C	Not enough data
01/22/81	A 022	X	C				No fringes found
01/22/81	B 022	X	C				Unreadable IDRS
01/29/81	B 029	X	C				Station problems
01/31/81	A 031	X	C				Missing IDRS
02/04/81	A 035	X	X	X	X	X	Data sent to BIH
02/04/81	B 035	X	C				Missing IDRS
02/11/81	A 042	X	X	X	X	X	Data sent to BIH
02/11/81	B 042	X	X	X	X	C	Not enough data
02/19/81	A 050	X	X	X	X	C	Not enough data
02/19/81	B 050	X	X	X	X	X	Data sent to BIH
03/01/81	A 060	X	C				No fringes found
03/01/81	B 060	X	X	X	X	X	Data sent to BIH
03/09/81	A 068	X	C				Station problems
03/09/81	B 068	X	X	X	X	C	Not enough data
03/16/81	B 075	X	C				No fringes found
03/17/81	A 076	X	C				Station problems
03/27/81	A 086	X	C				No fringes found
03/28/81	B 087	X	C				No fringes found
04/04/81	A 094	X	X	X	X		MFIT solutions
04/05/81	B 095	X	C				No fringes found
04/12/81	A 102	X	X	X	X		MFIT solutions
04/12/81	B 102	X	X	X	X		MFIT solutions
04/17/81	B 107	X	X				In backlog
04/18/81	A 108	X	C				No fringes found
04/26/81	A 116	X	X				In backlog
04/26/81	B 116	C					Antenna failure
05/03/81	A 123	X	C				Playback cancelled
05/03/81	B 123	X	X	X	X		S-band only
05/10/81	A 130	X	X	X	X		MFIT solutions
05/10/81	B 130	X	X	X	X		MFIT solutions
05/16/81	A 136	X	X	X	X		MFIT solutions
05/16/81	B 136	C					Station problems
05/24/81	A 144	C					Antenna failure
05/24/81	B 144	C					Antenna failure
05/31/81	A 151	X	X	X	X		MFIT solutions
05/31/81	B 151	X	X	X	X		MFIT solutions
06/08/81	A 159	X	C				IDRS bad
06/08/81	B 159	X	C				No fringes found
06/14/81	A 165	X	X	X	X		MFIT solutions
06/14/81	B 165	X	X	X	X		MFIT solutions
06/21/81	A 172	C					Pass cancelled
06/21/81	B 172	C					Pass cancelled
06/28/81	A 179	X	X	X	C		No delay data
06/28/81	B 179	X	X	X			Ready to MFIT
07/05/81	A 186	X	C				No fringes found
07/05/81	B 186	X	C				No fringes found
07/12/81	A 193	X	X	X	X		MFIT solutions
07/12/81	B 193	X	X	X	X		MFIT solutions
07/19/81	A 200	X	X	X	X		MFIT solutions
07/19/81	B 200	X	X	X	X		MFIT solutions
07/26/81	A 207	X	X	X	X		MFIT solutions
07/26/81	B 207	X	X	X	X		S-band only
08/01/81	A 213	X	X	X	X		MFIT solutions
08/02/81	B 214	X	X	X	X		MFIT solutions

Table 1 (contd)

Date	ID/DOY	Pass	Corr	PHASOR	FITDELAY	MASTERFIT	Notes
08/11/81	A 223	C					Pass cancelled ^a
08/12/81	B 224	C					Pass cancelled ^a
08/23/81	A 235	C					Pass cancelled ^a
08/23/81	B 235	C					Pass cancelled ^a
08/29/81	A 241	C					Pass cancelled ^a
08/29/81	B 241	C					Pass cancelled ^a

^aThese passes were cancelled to protect the DSS 14 antenna from further injury from the azimuth bearing problem

Table 2 Validated results for clock synchronization, Block I system

BLOCK 1 SYSTEM

=====

VALIDATED RESULTS FOR CLOCK SYNCHRONIZATION

=====

REPORT DATE 11/09/81 COVERING DEC THRU MAY 1981

----->		RELATIVE CLOCK PARAMETERS DSS43-DSS14							COMMENTS, NOTES, OR SIGNIFICANT EVENTS	RELATIVE CLOCK PARAMETERS DSS63-DSS14							<-----	
SET ID	NO. OBS. USED	REFERENCE TIME		CLOCK OFFSET		RATE X E-12		REFERENCE TIME		CLOCK OFFSET		RATE X E-12		NO. OBS USED	SET ID			
		DOY	HH MM SS	VALUE	ERR	VALUE	ERR	DOY		HH MM SS	VALUE	ERR	VALUE			ERR		
AC02	NS	337	*****	PASS CANCELLED	*****	*****	*****	NS=NOT SCHEDULED NG=NO BLK1 DATA *=BLK 1 BACKLOG %=UT1-PM PAIR -=CLOCK RESETS	336	*****	BAD IDRS DSS 14	*****	*****	NG	BC01			
AC08	8	343	05 25 41	-3 803	.016	- 08	.31		%	%	342	22 59 40	-1 976	.051	- 09	5.32	9	BC07
AC14	10	349	04 54.38	-3 529	.048	- 76	18		%	%	348	23 49 40	-1 463	.048	- 55	.66	8	BC13
AC23	9	358	08 34 43	-3.184	.061	23	20				358	*****	STATION PROBLEMS	*****	*****	*****	NG	BC23
AC30	NG	365	*****	STATION PROBLEMS	*****	*****	*****				366	05 42 40	-1 113	.133	.29	26	9	BC31
A122	NG	022	*****	NO FRINGES	*****	*****	*****				022	*****	NO FRINGES, NO TONES DSS14	*****	*****	*****	NG	B122
A130	11	030	07 04 59	-1 946	.062	- 08	18				029	*****	PASS CANCELLED	*****	*****	*****	NS	B129
A204	12	035	09 02 42	-1.809	.060	1 12	53				035	*****	INCOMPLETE IDRS	*****	*****	*****	NG	B204
A211	12	042	08 02 34	-1 550	.058	23	13		%	%	042	05 08 40	084	.070	.63	.49	9	B211
A219	6	050	07 04 59	-1.119	.130	- 08	.18		%	%	050	00 08 41	098	.205	83	28	8	B219
A301	NG	060	*****	NO FRINGES	*****	*****	*****				060	18 42 40	203	.073	- 55	27	8	B301
A309	NS	068	*****	PASS CANCELLED	*****	*****	*****				068	04 21 42	164	.093	89	80	8	B309
A316	NG	075	*****	NO FRINGES	*****	*****	*****				076	*****	PASS CANCELLED	*****	*****	*****	NS	B317
A327	NG	086	*****	NO FRINGES DSS 14 TONES WEAK	*****	*****	*****				087	*****	NO FRINGES, NO TONES DSS63	*****	*****	*****	NG	B328
A404	11	094	13 19 40	781	.071	2.83	44				095	*****	NO FRINGES, NO TONES DSS63	*****	*****	*****	NG	B405
A412	11	102	10 23 00	253	.066	1 01	66		%	%	102	03 09.30	- 842	.130	-1 22	1 52	7	B412
A418	NG	108	*****	NO FRINGES DSS 14 TONES WEAK	*****	*****	*****				107	06 44 36	488	.188	-2 26	0 55	7	B417
A426	9	116	12 08 34	1 545	.102	84	65				116	*****	DSS 14 ANTENNA FAILURE	*****	*****	*****	NS	B426
A503	NG	123	*****	DSS 14 IDRS LOST	*****	*****	*****				123	12 11.39	.426	.075	47	.19	9	B503
A510	13	130	06 39 39	2.188	.058	31	.99		%	%	130	03 54.41	403	.095	24	64	7	B510
A516	12	136	07 31 42	2 476	.046	13	.29				136	*****	BWS OSCILLATOR FAILURE	*****	*****	*****	NG	B516
A524	NS	144	*****	DSS 14 ANTENNA FAILURE	*****	*****	*****				144	*****	DSS 14 ANTENNA FAILURE	*****	*****	*****	NS	B524
A531	11	151	06 39.41	3 265	.081	- .19	32		%	%	151	03.48.39	618	.153	06	.22	7	B531

Table 2 (contd)

=====

VALIDATED RESULTS FOR CLOCK SYNCHRONIZATION

=====

REPORT DATE 11/09/81 COVERING JUN THRU AUG 1981

BLOCK 1 SYSTEM

[illegible]

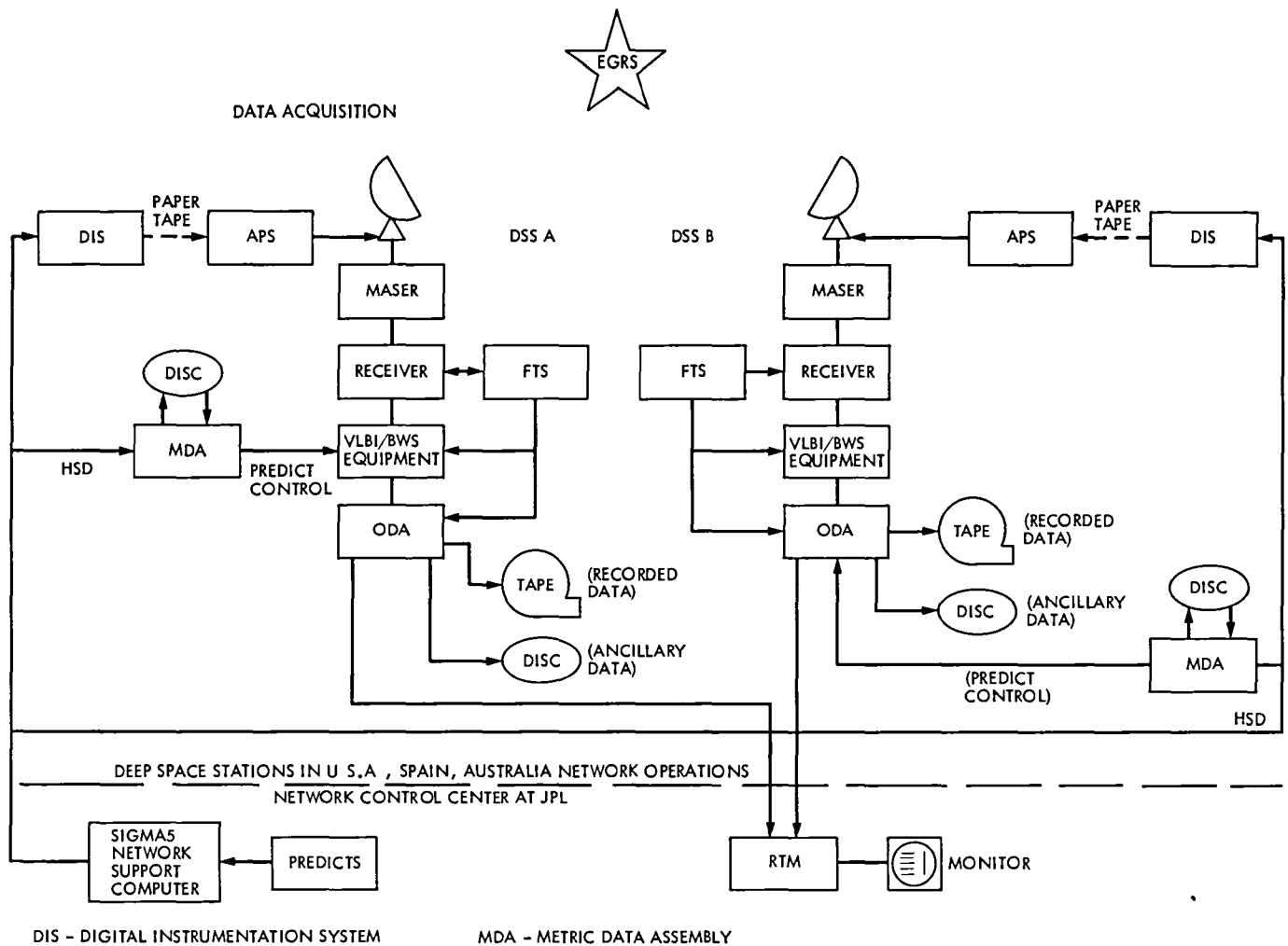


Fig. 1. VLBI system overview

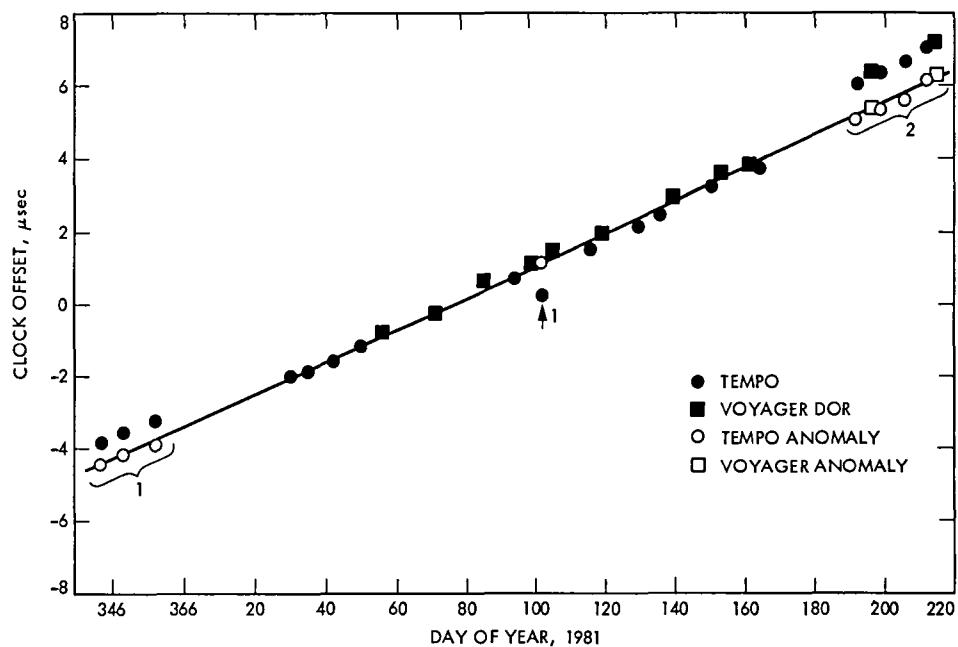


Fig 2 Clock offset, DSS 43/DSS 14

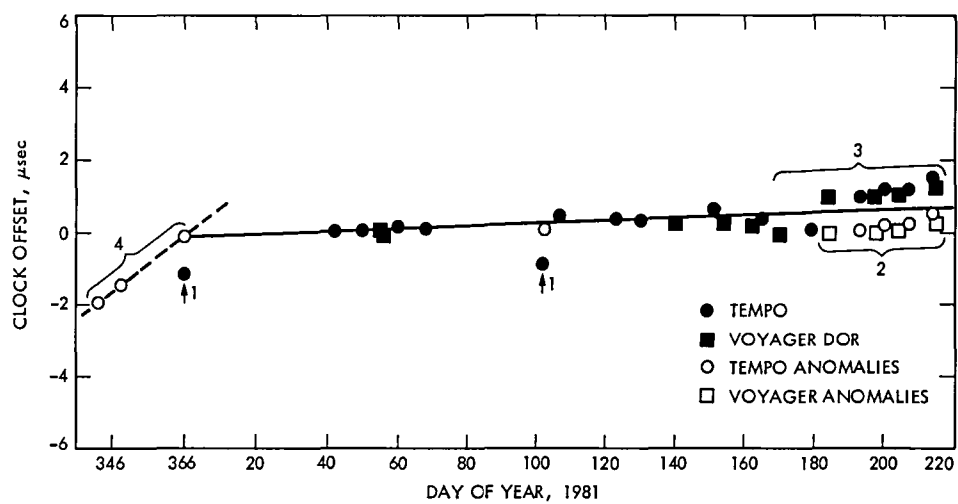


Fig 3 Clock offset, DSS 63/DSS 14

Synthesizer Stability Evaluation

E N Sosa and D A Tyner

Radio Frequency and Microwave Subsystems Section

An accurate and simple technique for measuring rms phase noise as a function of integration time for offset frequencies (10^{-3} Hz) close to the carrier is described. The accuracy of this method is verified by Allan Variance measurements in the time domain and through the use of Hadamard Variance for generating high-resolution spectral density measurements. Two different synthesizers are evaluated and the results are presented.

I. Introduction

Direct measurements of phase noise below 10 Hz from the carrier are not commonplace in industry. Special instrumentation and methods do exist for making close-in measurements well below 10 Hz, but these are costly as well as very time consuming. A recent need called for a simple but accurate technique that could be used by JPL and the synthesizer manufacturing industry. The need arose from a twofold requirement of the Phase 4 Receivers for VLBI. One aspect of the requirement called for cost-effective synthesizers because each receiver requires 12 synthesizers to provide a controlled selection of highly stable signals to be used for the second local oscillator signal. The other part of the requirement was related to an accurate in-house evaluation of close-in phase noise of a dozen potential candidate synthesizers. Thus, an accurate technique would assure the proper selection of the candidate synthesizer, and a simple test technique would be in keeping with lower cost units since this technique could be used for the acceptance testing of the production synthesizers. This report describes a method that is both accurate and amenable to production testing.

Typically, synthesizer manufacturers measure and state the stability of their units in terms of noise spectral density to

within 10 Hz of the carrier. System requirements, on the other hand, are stated in user terms, for example, phase error as a function of sampling or integration time. The test method described below measures the synthesizer stability as phase error in the time domain. Allan Variance measurements (in the time domain) were used to verify the phase error measurements. In addition, the same Allan Variance measurements were transformed to the frequency domain and thus characterized the synthesizer's noise spectral density to within 10^{-3} Hz from the carrier. The accuracy of the transformed data was verified via high-resolution spectral density measurements using the Hadamard Variance. The sections below describe the various measurement techniques, the time domain to frequency domain transformations, and the range of applicability.

II. Frequency Domain

A. Noise Spectral Density

To specify stability, synthesizer manufacturers' data sheets typically specify $\mathcal{L}(f_m)$, the ratio of the single sideband power of phase noise in a 1-Hz bandwidth f_m Hertz away from the carrier frequency to the total signal power. The usual range for f_m is from 10 to 10^6 Hz, with no data relating to the

extremely close-in phase noise, which is of interest to those concerned with medium-to-long-term stability

The customary method for generating $\mathcal{L}(f_m)$ curves utilizes spectrum analyzers, and the measurement techniques are straightforward and well-understood. To cover a wide range of f_m , several types of spectrum analyzers must be used and the typical measurement technique is shown in Fig 1. The different configurations (A and B) displayed in Fig 1 represent the method used to compensate for the limited offset range capability of typical spectrum analyzers. The newest generation of spectrum analyzers, like the HP 8568, provide most of the required range from offsets of 80 Hz to greater than 1 MHz away from the carrier, at the actual carrier frequency (Configuration A $f_o = 100$ MHz). Older analyzer models require more overlap with lower offset frequency domain analyzers such as the HP 3581 Wave Analyzer (max $f_o < 50$ kHz) or the HP 3585 low-frequency spectrum analyzer ($f_o < 40$ MHz). These instruments measure $f_m > 3$ Hz and 25 Hz respectively, but require the measurement setup of configuration B, which adds an external mixer prior to the low-frequency spectrum analyzer, and thus extends its upper-frequency limit. These techniques provide a direct measure of the noise power at different offset frequencies, and require the corrections delineated in the Appendix to derive the SSB noise in 1 Hz. Configurations A and B were used to generate the $\mathcal{L}(f_m)$ curve depicted in Fig 2 for the laboratory model synthesizer built by JPL, similar in design to the Mark III-Haystack Digitally Controlled Synthesizer (DCS). Spectral densities generated with the different instruments agree well where data overlap existed, and furthermore, data connect smoothly between different regions where measurements were not taken. The accuracy of the spectral density measurements is estimated to be within ± 3 dB. The other $\mathcal{L}(f_m)$ curve, also in Fig. 2, was obtained from the data sheet of a commercially available synthesizer. Both of these curves illustrate the frequency range that can be normally covered, but which does not get into the close-in range.

To extend the noise spectral density curves into the extremely close-in range, use was made of the Allan Variance method. Medium-to-long term stability performance in the time domain can be practically and meaningfully measured by fractional-frequency deviations (the square root of Allan Variance). The testing requires a statistical evaluation of the sampled synthesizer frequency, as it deviates from a reference standard. In general, fractional-frequency deviation does not provide a precise estimate of spectral density for offsets, f_m , far away from the carrier. However, for close-in phase noise, sufficient accuracy is obtainable, as shown later by the direct rms phase jitter measurements. Fractional-frequency deviation data were obtained for the two synthesizers whose noise spectral densities are given in Fig 2. The measurement technique is illustrated by the block diagram in Fig 3, with

the HP 5390 Frequency Stability Analyzer shown used to sample the beat note between the reference and synthesizer under test and to calculate the fractional-frequency deviation, $\sigma_y(\tau)$. Narrow noise bandwidths (f_h), corresponding to small beat notes (ν_b), were needed for measurement of close-in spectral density components of the continuous noise distributions. A wider f_h measurement is described in a later section to supplement this close-in data base with the middle to far-out offset frequency components. The measured stability performance for the two synthesizers is depicted in Fig 4 for $\nu_b = 1$ Hz. Medium-to-long-term stability performance is covered by the integration times, τ , from 10 to 10^3 seconds. Different noise processes in the synthesizers give rise to the slopes of the fractional-frequency deviation, $\sigma_y(\tau)$ plots in Fig 4. The noise process can be identified by its slope, and, by knowing the noise process, a time-to-frequency transformation can be made thus translating the accurate long-term stability measurements into close-in noise spectral density data.

B. Time-to-Frequency Transformations

Transformations from the time domain to the frequency domain have been described (Refs 1 and 2), and have been applied by Sward (Ref 3) to the highly stable hydrogen maser over the same offset frequency range of interest (10^{-3} to 10^6 Hz). A method similar to Sward's was used in this evaluation except that more exact formula (Ref 4) were used here for calculating the f_m corresponding to a particular integration time τ . A simple reciprocal conversion can lead to up to an order of magnitude error from the exact value. The more exact method involves calculating the f_m at the intersection points where the dominant noise process changes and gives rise to a different slope. For example, in Fig 4 the $\sigma_y(\tau)$ plot for the commercial synthesizer shows a break point (a change in slope) at $\tau = 40$ s. The simple reciprocal conversion leads to $f_m = 0.025$ Hz, while the more exact formula results in $f_m = 0.007$ Hz.

The extension of the noise spectral density plots were generated by approximating the experimental $\sigma_y(\tau)$ data with slopes corresponding to well-known noise processes. In the plots of Fig. 4, both units were considered to be composed of two noise processes. The commercial unit consisted of white frequency (slope = $-1/2$) and random walk frequency ($+1/2$), while the JPL-built unit exhibited flicker phase (slope = -1) and white frequency ($-1/2$).

The transformations used for the 3 different noise processes are

$$\mathcal{L}(f_m) = \frac{(\sigma_y(\tau) f_o \tau^{2.565})^2}{2.184 + \ln(f_h \tau)} f_m^{-1} \quad \text{flicker phase} \quad (1)$$

$$\mathcal{L}(f_m) = (\sigma_y(\tau) f_o \tau^{1/2})^2 f_m^{-2} \quad \text{white frequency} \quad (2)$$

$$\mathcal{L}(f_m) = (\sigma_y(\tau) f_o \tau^{1/2} 0.276)^2 f_m^{-4} \quad \text{random walk frequency} \quad (3)$$

in which f_o denotes the carrier frequency and f_h is the measurement bandwidth. The exact values for f_m were found by solving simultaneously for f_m and τ with pairs of the above equations whose corresponding noise processes intersect. The τ at the intersection point is known from the experimental fractional-frequency deviation data and the corresponding f_m is determined.

The transformed time domain data are shown plotted in Fig. 5 as extensions of the noise spectral density plots previously discussed. The transformed $\sigma_y(\tau)$ data “lines-up” well with the measured $\mathcal{L}(f_m)$ curves. However, to ascertain the accuracy of the transformations, close-in phase noise spectral density measurements were made with the HP5390 Frequency Stability Analyzer as configured in Fig. 3. This instrument also makes use of the Hadamard Variance, which was developed for making high-resolution spectral analysis. Frequency domain data is calculated from the time domain data provided by digital counters. The data obtained with the HP5390 is also plotted in Fig. 5, and support the simpler transformations used above.

III. Phase Noise in the Time Domain

It was previously mentioned that from a system's point of view, the desired stability parameter is phase noise in the time domain. In this section several techniques are described to measure phase noise or jitter as a function of integration time with conventional instrumentation. Also described below are how these measurements compare to the stability measured via the fractional-frequency deviation technique. Direct calorimetric phase jitter measurements were made and these were compared with the corresponding transformed $\sigma_y(\tau)$ measurements. Good agreement was obtained. Additional phase error data as a function of time was obtained by measuring temperature stability of several synthesizers. A further use of the phase jitter data involves the calculation of the relative power between noise and carrier in order to supply another frequency domain representation of synthesizer stability.

Using readily available instrumentation, a method was developed to measure phase error at selected integration times. The measurement block diagram for calorimetric true rms phase jitter is shown in Fig. 6. The technique involves heterodyning the phase noise spectra to baseband and calibrating the phase fluctuations to the measured rms voltage (Ref. 5). To generate phase jitter data as a function of

integration time, it was necessary to alternate true rms detectors. Each detector had a different (99%) response time and, with six detectors, integration times from 60 ms to 45 s were covered in six points. The integration time span is due to several different implementations of a true rms converter, utilized in the calorimetric detectors listed in Fig. 6. The shorter taus are obtained in high-speed monolithic circuits that can effectively integrate the high peak-to-rms-noise voltages into their equivalent root mean square values. Longer sampling times require more massive heat generating and detecting hardware and usually operate over much larger bandwidths. Sensitivity also degrades substantially for the longest τ detector (45 s) requiring a larger gain to measure the equivalent rms phase jitter (a dynamic range constraint for high spectral purity signal measurement).

Phase jitter is calculated from the measured raw noise and sinusoidal voltages by

$$\Delta\phi_{rms}(\tau) = \frac{(57\,296 \text{ deg/rad})}{\sqrt{2}} \frac{V_{noise}}{V_{beat}} \quad (4)$$

where $\Delta\phi_{rms}(\tau)$ is the rms jitter in degrees of a synthesizer at f_o , τ is the 99% response time of digital voltmeter (DVM), V_{noise} is the true rms voltage of the noise fluctuations, and V_{beat} is the true rms voltage of the calibration beat note. This assumes that the reference synthesizer is 10 dB more stable than the synthesizer under test. If the synthesizers are nearly the same, another $1/\sqrt{2}$ factor must be added to ascertain the individual synthesizer performance. The data obtained for the two synthesizers are plotted in Fig. 7 for a carrier frequency of 100 MHz.

The JPL-built synthesizer exhibited between 0.7 and 1.0 degrees rms of jitter for $\tau \leq 45$ s at 100 MHz. With the same setup, in the short integration times the commercial synthesizer manifested a hundredfold lower noise, but also drifted relatively more in the longer integration times. The fourfold degradation at $\tau = 45$ s, compared to the jitter for $\tau \leq 2$ s, demonstrates the type of phase drifts that can occur in some synthesizers, probably due to environmental factors.

It is interesting to note the nearly flat phase jitter slope at the low taus, for all of the curves depicted in Fig. 7. The dashed lines connecting the open symbols represent the corresponding phase jitter values obtained by applying a simple transformation of the $\sigma_y(\tau)$ data given for both synthesizers in Fig. 8. These fractional-frequency deviations, measured with $\nu_b = 10$ kHz and in a 25-kHz bandwidth (f_h), offer an excellent methodology comparison to the calorimetric data where $f_h = 15$ kHz. The transformation converts $\sigma_y(\tau)$ vs

τ to $\Delta\phi_{rms}$ vs τ by a straightforward scaling of the $\Delta f/f_o$ values. The equation used,

$$\Delta\phi_{rms}(\tau) = \sigma_y(\tau) \cdot f_o \cdot \tau \cdot 360 \quad (5)$$

also clarifies the observed short-term $1/\tau$ relation of the $\sigma_y(\tau)$ data and the independently measured constant $\Delta\phi_{rms}$ slope. Obviously as τ increases, the $\sigma_y(\tau)$ values must decrease equally in proportion to maintain a constant phase jitter slope. Thus a $1/\tau$ fractional-frequency deviation slope translates into a constant phase jitter. Furthermore, as the phase errors become larger for longer sampling times, the $\sigma_y(\tau)$ curves must flatten out and then eventually turn upward, typically beyond $\tau \cong 1000$ s (as observed in Fig. 8)

Equation 5 can also be applied to the $\sigma_y(\tau)$ data shown in Fig. 4, but due to the much smaller measurement bandwidth ($f_h = 1$ Hz) and $\nu_b = 10$ Hz, no meaningful relation exists to the measured calorimetric jitter taken in a 15-kHz bandwidth. The $\sigma_y(\tau)$ values of Fig. 4 would essentially relate to only the close-in noise spectral density as previously demonstrated, and thus exclude the predominant noise pedestal of the JPL-built DCS at offset frequencies between 0.1 and 10 kHz away from the carrier. An indication of this is found by comparing the $\sigma_y(\tau)$ values of Figs. 4 and 8, at $\tau = 1$ s. The JPL-built DCS suffers a 42-fold loss in stability, while the commercial unit degrades by only a factor of 3, for the increased measurement bandwidth from 1 Hz to 25 kHz, thus showing the importance of measurement bandwidth. Corroboration of the time and frequency domain data is enhanced by the observed downward trend of the transformed $\Delta\phi_{rms}$ of Fig. 7, for the short integration times of the commercial synthesizer. This type of behavior is expected due to the observed rapid drop in noise spectral density, for $f_m \geq 100$ Hz found in Fig. 5

The empirical data base established to generate the curves of Fig. 7 can also be used directly in the frequency domain, to generate another set of spectral densities. This is because the calorimetric method is basically a hybrid of the conventional $\mathcal{L}(f_m)$ measurement scheme and several variable τ detectors. The formula to generate $\mathcal{L}(f_m)$ is

$$\mathcal{L}(f_m) = 20 \log \left(\frac{V_{noise}}{\sqrt{2} V_{beat}} \right) - 10 \log(f_h) - 3 \text{ dB} \quad (6)$$

where V_{noise} is the true rms voltage at offset f_m Hz from f_o , V_{beat} is the true rms voltage of the $\nu_b = f_m$ beat note, f_m is the wave analyzer's tuned frequency, f_h is the noise bandwidth, and the (-6 dB or -3 dB) factor is for SSB noise with equivalent sources or for a reference with 10-dB lower noise

This measurement also utilizes the test configuration in Fig. 6 (see Ref. 6), and in Table 1 typical phase jitter is listed for each synthesizer tested and their related calculated spectral densities. The offset frequency, f_m , in this case was a distributed parameter due to the wideband low-pass filter used (Fig. 6 $f_h = 15$ kHz) and thus the calculated value represents an integration of the noise sidebands normalized to a 1-Hz bandwidth without a specific f_m (i.e., $DC < f_m < 15$ kHz). Even so, the calculated results match well with the measured $\mathcal{L}(f_m)$ displayed in Table 1 or in Fig. 5. Additional jitter data were taken as a function of f_m by utilizing the wave analyzer as a variable tuned filter, as shown beneath the dotted line in Fig. 6. Results taken in this manner demonstrated the relationship of wider resolution bandwidths and variable offset frequency f_m , and confirmed the curves of Fig. 2. It is thus practical to measure spectral densities of noise sidebands by the apparatus used to measure phase jitter as a function of integration time.

IV. Temperature Stability

Phase stability as a function of time also becomes temperature related, due to the susceptibility of synthesizers to environmental changes. The temperature coefficient of output phase was measured and can be related to long-term phase instabilities. The measurement block diagram used is found in Fig. 9. The data were taken by applying a temperature step ($\Delta T = 20^\circ\text{C}$) to the synthesizers under test and watching their phase change with respect to the reference synthesizer, at ambient (Ref. 7). An exponential phase response resulted from the applied step temperature change, and the phase stabilized at a new value after 20 or 30 min. By dividing the phase step by the temperature step, one arrives at a number that indicates the susceptibility of phase changes in the expected receiver environment. The equation is simply

$$\frac{\Delta\phi}{\Delta T} = \frac{\phi_{init}(T = T_1) - \phi_{final}(T = T_2)}{T_2 - T_1} \quad (7)$$

where $\Delta\phi/\Delta T$ is the change in output phase per degree centigrade change in temperature, ϕ_{init} and ϕ_{final} are steady-state values displayed on the vector voltmeter before and after the applied thermal step $\Delta T = T_2 - T_1$.

Results of this test are tabulated in Table 2 for each synthesizer and these data indicate that a synthesizer such as the JPL-built DCS, with a lower parts count, is about 5 times less sensitive to environmental variations than the commercial unit tested.

Another attribute of the temperature stability measurement scheme is the ability to generate a frequency stability value,

$\Delta f/f_o$, for the equivalent thermal time constant of the synthesizer tested. The same exponential phase vs time curve discussed above is used to calculate the $\Delta f/f_o (\tau_{therm})$. The $1/e$ (or 63% of the final phase value) time interval, and the actual $\Delta\phi$, has been normalized to the operation frequency and recorded in Table 2. These numbers represent nonstatistical phase drift susceptibilities found at longer integration times ($\tau > 1$ min), and thus further buttress the long-term stability analysis. The expression used to compute the temperature stability displayed in Table 2 and Fig 8 is

$$\frac{\Delta f}{f_o} (\tau_{therm}) = \frac{\phi_{init} (t=0) - \phi_{63\%} (t=\tau_{therm})}{360 \cdot \tau_{therm} \cdot f_o} \quad (7)$$

where $\phi_{init} (t=0)$ is the initial steady phase, prior to applied step ΔT , $\phi_{63\%} (t=\tau_{therm})$ is the phase shift incurred after τ_{therm} seconds, and τ_{therm} is 63% or $\phi_{final} (t \geq 3 \tau_{therm})$ measured in seconds

The long-term $\sigma_y (\tau)$ values seem to connect fairly well to these indirectly generated $\Delta f/f_o$'s as the expected up turn in the two curves occurs at their respective thermal time constants

V. Conclusion

In the absence of conventional instrumentation for measuring close-in noise, a method has been developed that not only utilizes readily available equipment, but that can be used for production testing. The measurement technique utilizes the different time constants inherent in the different types of true rms meters. Two different synthesizers were measured, and the results were in very good agreement with stability data obtained by more sophisticated methods and instruments. Allan Variance was used to establish that, via transforms, fractional-frequency deviations gave a good representation of close-in phase noise. This was also supported by spectral

density measurements using the HP5390 Frequency Stability Analyzer. Allan Variance measurements were subsequently transformed to phase noise as a function of time and compared to the phase jitter measurements made by the method herein described. Good agreement was obtained up to $\tau = 45$ seconds, which corresponds to 6×10^{-3} Hz. Knowledge of the noise process for the particular synthesizer determines the slope beyond 45 seconds. More accurate determination of $\Delta\phi_{rms}(\tau)$ beyond 45 seconds can still be accomplished by $\sigma_y(\tau)$ measurements except that these take longer test time. Other production-oriented measurement techniques that offer potential are digital spectrum analyzers that utilize the FFT (Fast Fourier Transform) to transform time domain samples in parallel into the frequency domain with a minimal delay, or a new class of (digital) video chart recorders that can provide statistical averaging of the phase drifts as a function of integration time. A few major synthesizer vendors currently utilize the HP 3582 FFT Spectrum Analyzer for their production testing of close-in ($f_m > 1$ Hz) noise spectral densities for their top line, low-noise synthesizers. The instrumentation just mentioned is capable of measuring considerably closer to the carrier, within 10^{-2} Hz, but little need typically has been expressed for such information. It turns out that VLBI receivers also require measurements of phase error as a function of integration time, and any frequency domain method would necessitate accurate transformations from the frequency to the time domain to assure compliance with the overall system design.

While the results of the measurement techniques herein described agreed well with transformed data, the overall limits of applicability were not established. Evidently the measurement technique is valid for the case of synthesizers locked to a spectrally pure hydrogen reference source, and for the integration times applied. Further investigation into the nonstationary (random walk) noise phenomenon may establish limits to the application, and may offer the optimal method of measuring long-term instabilities of phase.

References

1. Scherer, D , "Design Principles and Test Methods for Low Phase Noise RF and Microwave Sources," Hewlett-Packard RF and Microwave Measurement Symposium and Exhibition, June 23-24, 1980 (an update to his earlier report in *Microwaves*, pp 113-122, April and May 1979)
2. Rutman, J , "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources Fifteen Years of Progress," *Proc of IEEE*, Vol 66, No 9, pp 1048-1075, Sept 1978.
3. Sward, A , "Measurement of the Power Spectral Density of Phase of the Hydrogen Maser," *Quarterly Technical Review*, Vol 1, No 4, pp 30-33, Jet Propulsion Laboratory, Pasadena, Ca , Jan 1972.
4. Burgoon, R , and Fisher, M C , "Conversion Between Time & Frequency Domain of Intersection Points of Slopes of Various Noise Processes," *32nd Annual Frequency Control Symposium Proceedings*, National Technical Information Service, pp 514-519, 1978
5. Walls, F L., Stein, S R , Gray, E , and Glaze, D J , "Design Consideration in State-of-the-Art Signal Processing and Phase Noise Measurement Systems," *30th Annual Frequency Control Symposium Proceedings*, National Technical Information Service, pp 269-274, 1976
6. Fischer, M C , "Frequency Domain Measurement Systems," *10th Precision Time and Time Interval Applications and Planning Meeting Proceedings*, December 1978 Technical Information & Administrative Support Division, Goddard Space Flight Center, Greenbelt, Md (Also available in *Microwaves*, July 1979, pp 66-75, entitled "Analyze Noise Spectrum With Tailored Test Gear")
7. MacConnell, J., and Meyer, R , "L-Band Frequency Multipliers Phase Noise Stability and Group Delay," *Deep Space Network Progress Report, TR 32-1526*, Vol X, pp 104-109, Jet Propulsion Laboratory, Pasadena, Ca , Aug 1972

Table 1. Calculated average noise spectral density derived from phase jitter data

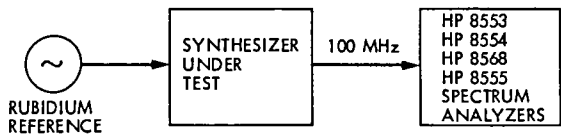
Synthesizer Model	Actual $\Delta\phi$ rms, deg rms	Calculated $\mathcal{L}(f_m)^a$ dBc/Hz	Measured $\mathcal{L}(100\text{ Hz})$, dBc/Hz
JPL-built DCS	0.81	-82	-75
Commercial unit	0.0083	-122	-131

^a f_m , the offset frequency, is now a distributed parameter, somewhere between 1 and 15,000 Hz

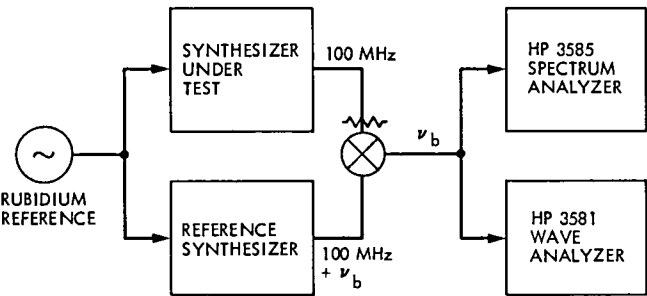
Table 2. Temperature coefficient of output phase and frequency deviation test results

Synthesizer	$\Delta\phi/^\circ\text{C}$, deg peak/ $^\circ\text{C}$	$\Delta f/f_o$ (τ_{therm})	τ_{therm} , s
JPL-built DCS (variation of 3)	2.2 ± 1	$(5.6 \pm 4) \times 10^{-13}$	720 ± 300
Commercial unit	10.2 ± 1	5.1×10^{-13}	3000 ± 100

CONFIGURATION A



CONFIGURATION B



KEY PARAMETERS

ANALOG SPECTRUM ANALYZERS	MIN OFFSET ^a f_m , Hz	CONFIGURATION
HP 8568	80	A
HP 8555	800	A
HP 8554	800	A
HP 8553	100	A
HP 3585	25	B

WAVE ANALYZER^b

HP 3581	30	B
---------	----	---

^aWHERE $\mathcal{L}(f_m) < -80$ dBc/Hz AND $\nu_b = f_m$

^bMAX OFFSET LIMITED HERE TO $f_m \leq 50$ kHz

Fig. 1. Noise spectral density, $\mathcal{L}(f_m)$, measurement block diagram

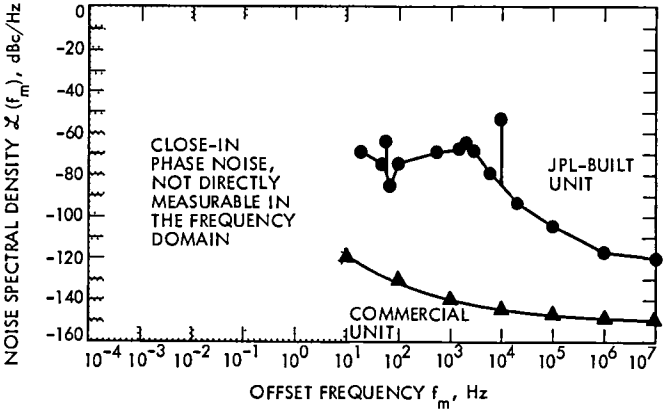
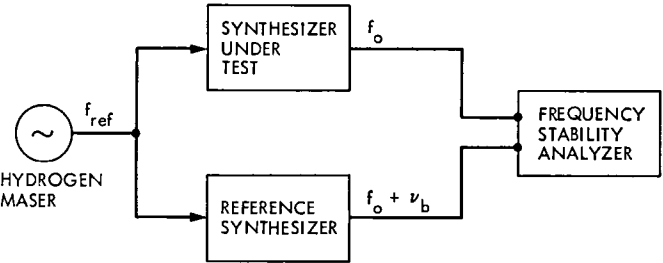


Fig. 2. Measured SSB phase noise of JPL-built DCS and commercial unit



KEY PARAMETERS

FREQUENCY STABILITY ANALYZER	HP 5390
$\sigma_y(\tau)$ (TIME DOMAIN)	
BEAT NOTE AND NOISE BANDWIDTH (f_h)	
$\nu_b = 10$ Hz	$f_h = 10$ Hz
$\nu_b = 10$ kHz	$f_h = 25$ kHz
REFERENCE SYNTHESIZER	
HP 8662 (OR)	$f_o = 100 + \nu_b$ MHz
HYDROGEN MASER	$f_o = 100$ MHz
EXTERNAL REFERENCE	
HYDROGEN MASER	$f_{ref} = 5$ MHz
NUMBER OF SAMPLES	$n \geq 25$
SYNTHESIZER UNDER TEST	
JPL-BUILT DCS	$f_o = 100$ MHz
COMMERCIAL UNIT	$f_o = 100$ MHz
TYPICAL NOISE FLOOR	$\sigma_y(1s) < 1 \times 10^{-14}$

Fig. 3. Fractional frequency deviation measurement block diagram

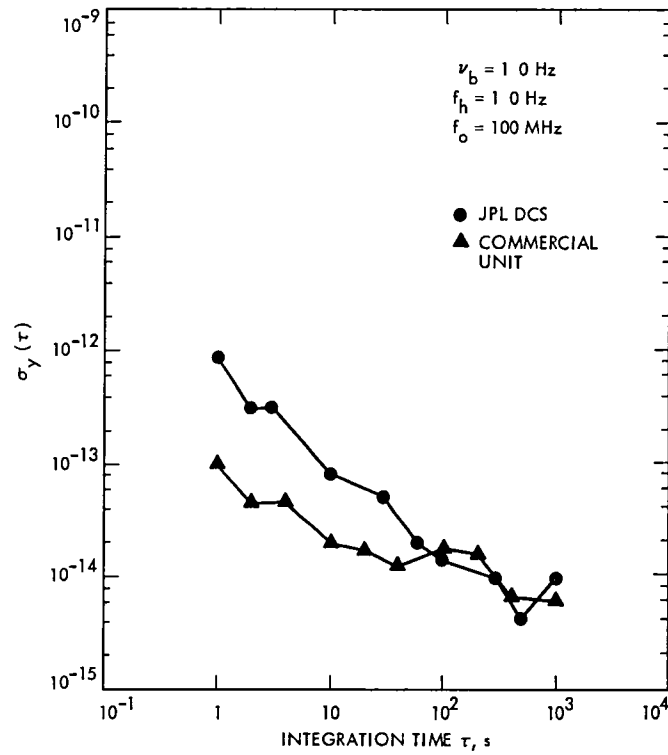


Fig 4 Fractional frequency deviation of JPL-built DCS and commercial unit

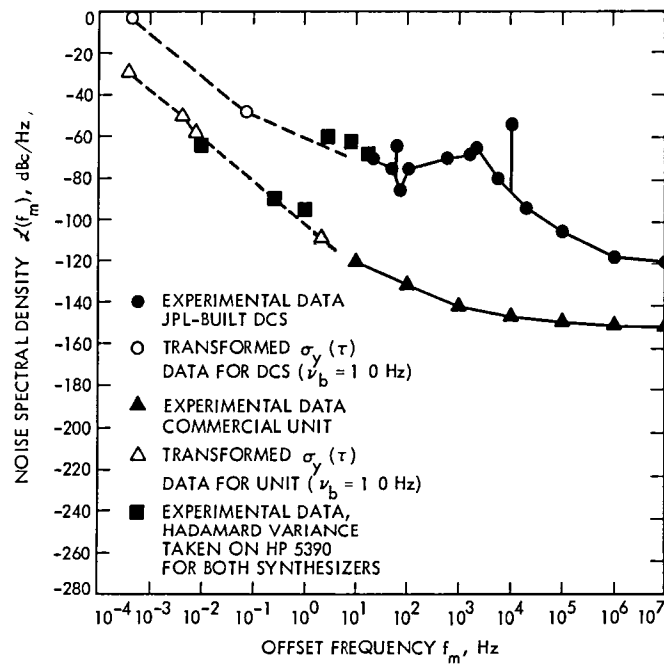
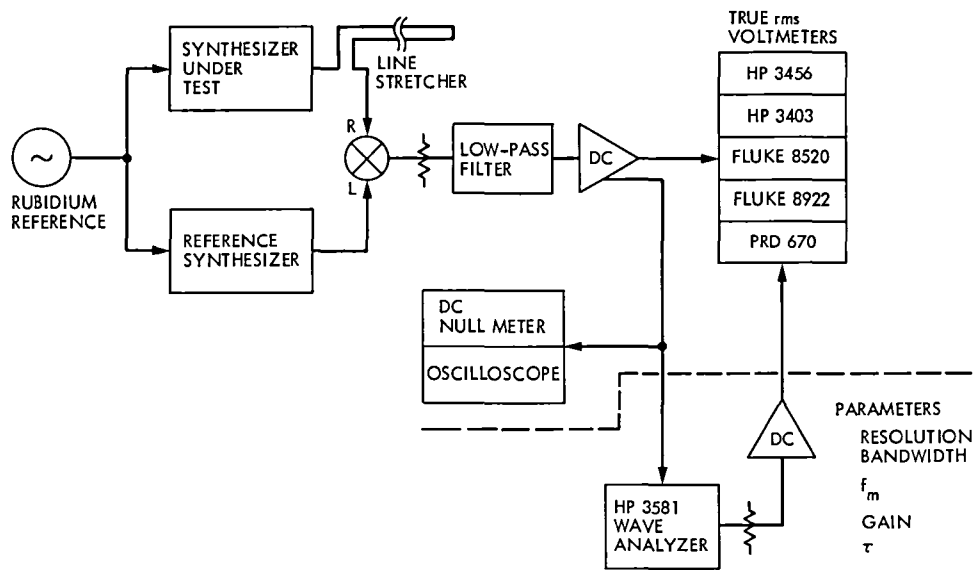


Fig 5 Composite noise spectral density



KEY PARAMETERS

TRUE rms VOLTMETER INTEGRATION TIMES (τ)

HP 3456	0.06 s
FLUKE 8520	0.10 s
HP 3403 (FAST/SLOW)	1.0/10 s
FLUKE 8922 (FAST/SLOW)	2.0/7 s
PRD 670 (DC RESPONDING)	45.0 s

MEASUREMENT BANDWIDTH

DYNAMICS 6050 DC AMP	DC TO 15 kHz
WAVE ANALYZER (HP 3581)	
$\Delta\phi_{rms}$ vs BW	3 Hz \leq RES BW \leq 300 Hz
$\Delta\phi_{rms}$ vs f_m	50 Hz $\leq f_m \leq$ 5 kHz

TYPICAL NOISE FLOOR

\mathcal{L} (1 kHz) < -130 dBc/Hz
$\Delta\phi_{rms}$ < 0.001 deg rms (in 15 kHz)

Fig 6 Calorimetric phase jitter vs τ measurement block diagram

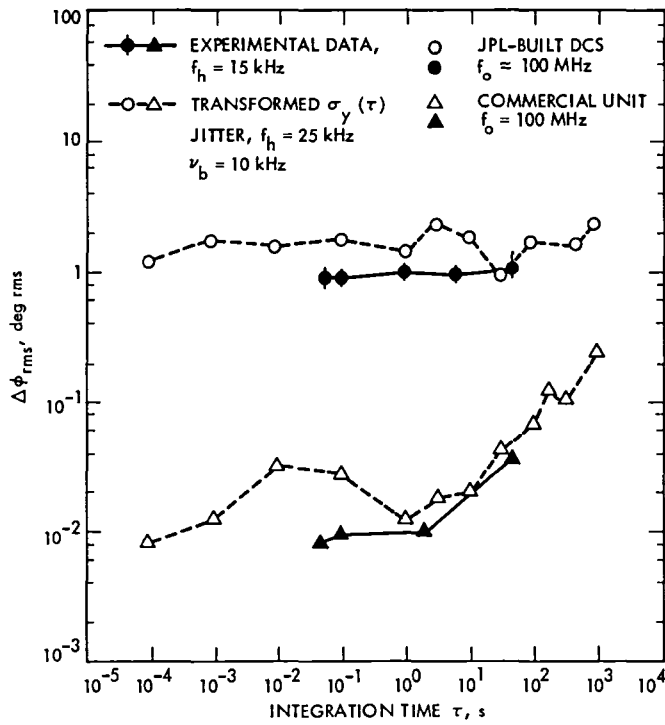
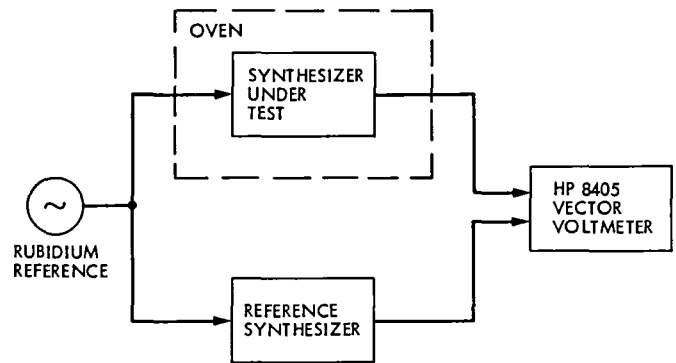


Fig. 7 Rms phase jitter vs τ of JPL-built DCS and commercial unit



KEY PARAMETERS

PHASE SHIFT RESOLUTION	0.1 deg UP TO 1 GHz
PHASE SHIFT ACCURACY	± 1.5 deg
TEMPERATURE ACCURACY	$\pm 0.2^\circ\text{C}$
TEMPERATURE STEP, ΔT	20°C
TYPICAL NOISE FLOOR	
TEMPERATURE COEFFICIENT OF OUTPUT PHASE	$< 0.05 \text{ deg}/^\circ\text{C}$
$\Delta f/f_o$ (WHERE $f = 100 \text{ MHz}$ $\tau \approx 60 \text{ s}$)	$< 1 \times 10^{-14}$

Fig. 9. Temperature coefficient of output phase measurement block diagram

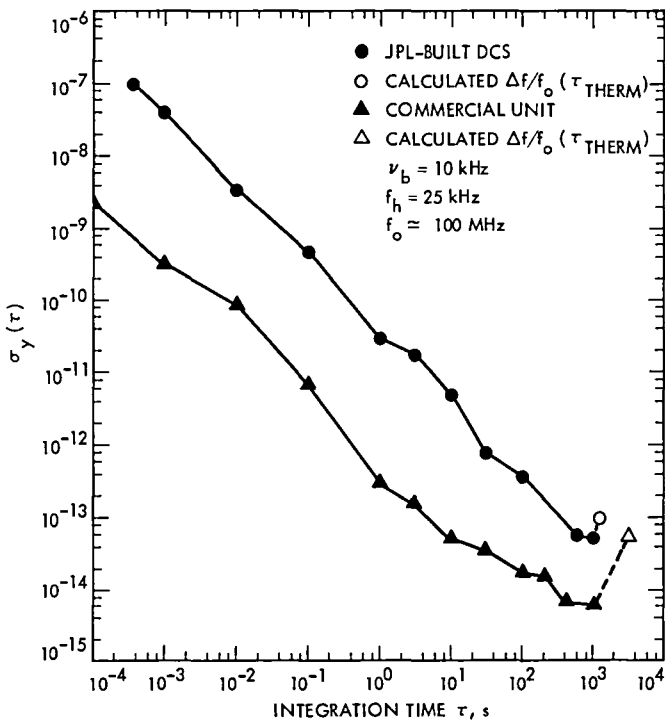


Fig. 8 Fractional frequency deviation of commercial unit and of JPL-built DCS

Appendix

Power Spectral Density Calculations

I. Spectrum Analyzer (Fig. 1)

$$\mathcal{L}(f_m) = -(P_{carrier} - (P_{noise} - 6 \text{ dB} + 2.5 \text{ dB} - 10 \log(f_h) - 3 \text{ dB}))$$

where

$\mathcal{L}(f_m)$ is in dBc/Hz, or dB below f_o the carrier, measured at the offset f_m , in 1-Hz equivalent noise bandwidth

$P_{carrier}, P_{noise}$ are measured powers in dBm

-6 dB is a factor for SSB and the displayed carrier power in rms, rather than peak power

+2.5 dB corrects for log amp at IF

f_h is noise bandwidth, $\cong 1.15 \times$ resolution BW

(-3 dB or -0 dB) is for two equivalent sources or if reference is 10-dB lower noise

II. Wave Analyzer (Fig. 1, Configuration B; Fig. 6)

$$\mathcal{L}(f_m) = -(P_{carrier} - (P_{noise} - 6 \text{ dB} + 1.05 \text{ dB} - 10 \log(f_h) - 3 \text{ dB}))$$

where

+1.05 dB is the correction factor for white noise as average responding detector is calibrated to read true rms, for a discrete signal

f_h is equivalent noise bandwidth, $\cong 1.10 \times$ resolution BW, as typical wave analyzers utilize narrower IF filters

Investigation of Digitally Controlled Oscillator Operability Problems

W. D. Weisman

Control Center Operations Section

In order to provide adequate levels of flight project support, the elimination of procedural errors by Deep Space Network personnel is an important consideration. This article describes an investigation designed to identify and correct procedural errors involving digitally controlled oscillator operations.

I. Introduction

The digitally controlled oscillator (DCO) is a microprocessor-based exciter frequency controller that serves as the prime uplink control unit on the 64-meter and 34-meter subnets. The DCO is controlled by a Deep Space Station (DSS) operator through either the Metric Data Assembly (MDA) in automatic mode, or a dedicated terminal in manual mode.

The DCO has been the source of numerous support problems since its implementation in July 1980. These problems can be categorized as either hardware, software, or operational problems. The identification and correction of operational problems is the responsibility of the Operations Control Group, Control Center Operations Section. These problems are documented, through the DSN Discrepancy Reporting System, as procedural discrepancy reports (DR's). The large number of procedural DR's involving the DCO indicate that the DCO suffers from poor operability. The term "operability" refers to the relative ease or difficulty encountered by an operator in attempting to operate a system correctly.

The methods used for identifying and correcting hardware and software problems are fundamentally different from the

methods used to identify and correct operability problems. Hardware and software problems can be dealt with through the use of step-by-step troubleshooting procedures, which lead directly to identification of a bad component or an incorrect line of code. Operability problems, however, are less tangible in nature because they originate at the interface between the DCO and its human operator. This person/machine interface is therefore the area that must be investigated when attempting to identify and correct operability problems. The methods used to investigate operability of a system are provided by the fields of behavioral science and human engineering.

II. Investigation

A. Investigative Strategy

Operability in a complex system is the result of the interaction of many variables. These variables provided a framework for analysis of the data. The variables deemed most relevant to DCO operability were.

- (1) The person/machine interface
 - (a) Physical operations environment.

- (b) Complexity and comprehension of input and output
- (c) Time constraints
- (d) Personality factors
- (2) Information transfer procedures
 - (a) Information transfer medium (voice, hard copy, CRT).
 - (b) Memorization requirements
- (3) Procedural documentation
 - (a) Metric Data Assembly Software Operator's Manual (MDA-SOM).
 - (b) Other documentation (Synthesizer Controller Technical Manual, DCO Procedures, TWX's)

B. Data

The data for this investigation were extracted from 45 procedural DR's involving the DCO over a period of 14 months from August 1980 to September 1981. Additional information regarding DCO operability was provided by interviews with station operators, reviews of DCO procedural documentation, and observations of DCO operations at DSS 14. Analysis of the data attempted to answer three major questions

- (1) Were procedural errors distributed evenly throughout all DCO-equipped stations?
- (2) Did the procedural error rate for DCO operations remain steady over time, or did it fluctuate in response to external events?
- (3) What types of errors occurred most often?

C. Results

No evidence was found to indicate that any DCO-equipped station was more prone to operability problems than any other DCO-equipped station. Within the 14-month period, the mean number of procedural DR's per station was 6.4, with a standard deviation of 2.87.

Several interesting phenomena were evident when the DR's were examined over time (see Fig. 1). The gradually decreasing error rate seen between August 1980 and February 1981 indicated that learning and familiarity over time aided station operators in reducing the procedural error rate to near zero. This gradually decreasing error rate is what one would expect to see following introduction of a new system. The dramatic increase in procedural errors seen in March, April, and May of 1981 was due to the implementation of new MDA software, which the station operators also needed time to become accustomed to.

The procedural DR's were also categorized by type of error. The following list indicates the most prevalent types of errors.

- (1) Incorrect command
- (2) Incorrect command syntax
- (3) Commands entered in wrong sequence
- (4) Wrong predicts selected.
- (5) Incorrect frequency entered

The last category, incorrect frequency entered, accounted for more than 25% of all procedural errors involving the DCO. No other problem category accounted for more than 10% of the total errors.

Interviews with station operators and observations of DCO operations pointed out several contributory factors to poor operability. Many errors occurred during high-activity periods at the station, especially during tune-in shortly after acquiring the downlink. At least two procedural errors also occurred during shift changes. Another factor is the time constraint associated with several DCO commands, which must be entered at least 4 minutes prior to the start of tuning.

A review of procedural documentation indicated that complete procedures did not exist for all possible modes of operation and that the MDA-SOM did not contain adequate uplink tuning procedures. Station operators also noted that procedural documentation was often not timely and not formatted so as to be easily usable.

Most of the problems regarding procedural documentation have recently been alleviated, as described below.

D. Discussion

The investigation results indicated that entering frequencies was the most error-prone aspect of DCO operations. These frequencies are large numbers, typically containing from 9 to 14 digits. The types of errors occurring most often involve transposition of digits or substitution of incorrect digits for correct digits. These errors are due to the memorization requirements of information transfer procedures, especially when information is being transferred via voice net. The capacity of human short-term memory is generally limited to about 7 discrete items of information. These items can be kept active indefinitely in memory by internal rehearsal, however, the potential for incorrect memorization increases as the number of items to be remembered increases. Multidigit numbers, therefore, become progressively harder to remember correctly as the number of digits is increased.

Another important feature of human memory is that even visually presented information is stored in acoustic form. For example, if a person is given a list of letters of the alphabet to memorize, and makes an error in recalling the letter "F", he is much more likely to report it as an "X" than an "E". Although "F" and "E" share common visual features, "F" and "X" share a common initial sound. This phenomenon is known as acoustic confusion. When memorizing large numbers, therefore, 2's are confused with 3's, 4's with 5's, and 6's with 7's.

Based on these findings, recommendations were made in June 1981 to carefully double-check frequencies, especially when they are passed to the station operator via voice net. Since this recommendation was implemented, only one procedural error involving entering a wrong frequency into the DCO has occurred.

Another possible source of error in DCO operations is the nonstandardization of delimiters. For example, certain MDA

commands for entering the date and time use spaces as delimiters, while equivalent DCO commands use colons as delimiters. This lack of standardization was the cause of at least one procedural error.

Procedural documentation was often mentioned by station operators as contributing to poor operability. Procedural documentation must accommodate both heuristic thinkers, who reason in global or system terms, and analytic thinkers, who reason in linear terms.

In May 1981 a special team was formed to rewrite the then-current DCO operations procedures. These new procedures were implemented in early June and were largely responsible for the reduction in procedural errors since then. As of this writing, only one procedural DR involving the DCO has been written in October.

Bibliography

- 1 Lindsay, Peter H , and Norman, Donald A., *Human Information Processing*, Academic Press, New York, 1977.
- 2 Schneiderman, B , *Software Psychology Human Factors in Computer and Information Systems*, Winthrop, Cambridge MA, 1980
- 3 Tausworthe, Robert C , *Standardized Development of Computer Software Part II – Standards*, JPL Publication SP 43-29, Jet Propulsion Laboratory, Pasadena, Calif , 1978
4. Wackley, J. A , "The DSN Tracking System", in *Telecommunications and Data Acquisition Progress Report 42-63*, Jet Propulsion Laboratory, Pasadena, Calif., March and April 1981, June 15, 1981.

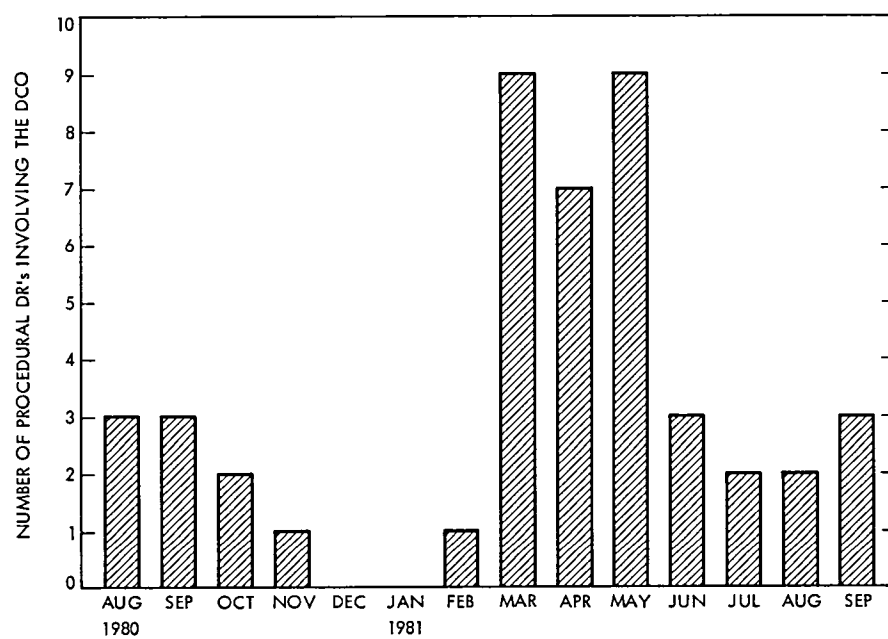


Fig 1. Procedural DR's per month since DCO installation

Emergency Telemetry Performance

W Tisdale
Deep Space Network Support

Often an emergency telemetry mode is automatically selected by a spacecraft during emergencies. During such emergencies the DSN must acquire the spacecraft telemetry as rapidly as possible under weak fluctuating signal conditions. Test data herein show the bit error properties of one emergency telemetry mode candidate with a very low bit rate of 4 BPS. Analysis reveals the E_b/N_0 necessary to obtain the performance BER threshold and associated acquisition times.

I. Introduction

Receiving telemetry during a spacecraft emergency is far from optimal. The spacecraft may be experiencing an RF power failure or loss of antenna point under dynamic conditions to produce weak intermittent downlink RF and telemetry signals.

Such a problem was presented in support of the NASA ISPM mission design. The search for a low bit rate to increase the chances of receiving emergency telemetry began, and telemetry tests were performed to demonstrate the capability of the DSN to support 3.93 BPS, coded and uncoded. Coding options include Reed-Solomon (255,223) outer code concatenated with a convolutional (7, 1/2, 3) inner code. The uncoded case is a contender since acquisition is more rapid.

The objectives of the tests reported herein are to examine the BER performance and signal acquisition time for both coded and uncoded links around the specified BER threshold of 10^{-3} . The input E_b/N_0 needed for the threshold is to be established.

II. Test Configuration and Test Techniques

All tests were performed in Compatibility Test Area 21 (CTA-21) at JPL. A block diagram of the test setup is depicted in Fig 1. RF and telemetry subsystems include SCA, an S-band transmitter, a Block-III RF receiver, a Block-III SDA, SSA and MCD. The test configuration included the convolutional coding for portions of the tests and uncoded for the remaining. CTA-21 had no capability for Reed-Solomon coding-decoding, therefore, no concatenated code tests were performed. Modulation included biphase modulation of symbols upon a 40 kHz squarewave subcarrier. Then this spectrum is phase modulated at 37 degrees mod index upon the S-Band carrier. All SNR levels were set by the standard Y-Factor technique.

Two problems impede progress when running telemetry tests at very low bit rates and with efficient codes. The first is test time. Sufficient test time is needed to accrue enough bit errors (N_e) for a high confidence interval and accuracy. With a given bit error rate BER (P_e) and bit rate (R), test time may be calculated by the following relation

$$T_e = \frac{N_e/P_e}{R/60} \text{ minutes}$$

To accrue 100 errors at a $P_e = 1 \times 10^{-3}$ takes 6 hours and 11 minutes at 4.49 BPS. This length of time represents a large expenditure of manpower and test area time. The difficulty in making test setups is compounded by the use of the efficient convolutional code. Efficient codes result in a very "steep" BER curve. If the E_b/N_0 is set too high, an insufficient number of errors occur and the BER result is inaccurate. If the E_b/N_0 is set too low the MCD will lose node synchronization.

When telemetry tests are performed at CTA-21, burst errors accumulate within high speed blocks. There are 960 bits per block. If the BER is plotted as the bit errors occur, as in Fig. 2, the BER will start high and approach the correct BER asymptotically. Not until about 108 errors (or 7 bursts averaging 15^+ errors) have occurred does the BER approach within 30% (or 0.1 dB) of the asymptotic BER value. Further work is needed to develop this technique analytically and empirically. For the testing analysis, a goal of 100 errors was established in the planning phase but due to the logistics of scheduling the test area and making long test runs, this was not found to be practical.

Testing for the uncoded 4.0 BPS proceeded uneventfully except for the large amount of test time needed.

III. Interpretation of Test Data

Test data in Table 1 is presented graphically in Fig. 3. The convolutionally encoded telemetry test results are compared against Linkabit baseband simulation data, also plotted in Fig. 3, to obtain total modulation-demodulation loss. At the performance threshold of $P_e = 10^{-3}$ the total system loss is 4.0 dB. This is equivalent to a baseband performance of $E_b/N_0 = 3.0$ dB. Note that with Reed-Solomon concatenation the system would be essentially errorless at 3.0 dB.

For input E_b/N_0 levels less than 7.0 dB, the BER performance is erratic. Erratic performance is qualitatively described as momentary and/or permanent data inversion, loss of MCD node synchronization, and very large data bursts beyond the ability of the Reed-Solomon concatenation to correct.

The uncoded test results in Fig. 3 are extrapolated down to a BER of 10^{-3} , showing that an input E_b/N_0 of 9.2 dB is necessary to obtain threshold performance. Tests 1, 2, and 3 (of Table 1) experienced 59, 7, and 128 bit errors respectively. To estimate the range of accuracy for these tests about 100 bit errors are needed to obtain a 90% confidence interval of being

within ± 0.1 dB. Approximately 10 points are needed for an 80% confidence to be within ± 0.5 dB (Ref. 1). No adverse behavior of the uncoded BER performance was apparent even though the RF carrier tracking margin ranged from 3.6 dB to 5.6 dB, which is well into the cycle slip region. Uncoded system loss is shown to be about 2.4 dB when the BER test data are compared to the theoretical curve.

IV. Acquisition Behavior

Due to the nature of the spacecraft emergency mode, quick acquisition time may be more desirable than a low bit error rate. If the spacecraft has a weak fluctuating SNR level that occasionally exceeds threshold, the acquisition time should be short by comparison to the time above threshold.

The RF and telemetry string acquisition-time calculations are broken down by subsystems and recorded in Table 2. Notable is the large percentage of time needed for Block III SDA to acquire. Also, the Reed-Solomon frame sync requires a very large percentage of the total acquisition time.

RF acquisition times of 5 and 10 minutes were assumed for E_b/N_0 levels of 7.0 and 9.0 dB, respectively. SDA, SSA, and MCD acquisition times were calculated from the instructions given in the DSN-Flight Project Interface Design Handbook. Reed-Solomon acquisition time is based on acquiring a 10080-bit frame within a range between 1/2 frame to 4 frames at 3.93 BPS.

V. Summary

Table 3 summarizes the BER and acquisition performance with use of three coding schemes: uncoded, convolutional, and Reed-Solomon convolutional concatenation. The uncoded telemetry needs a 9.2 dB input E_b/N_0 level for threshold performance but has the least acquisition time: 28 minutes and 18 seconds. The next best performer is the convolutionally code telemetry, which requires an input E_b/N_0 of only 7.0 dB, but the acquisition time has increased to 1 hour and 49 minutes. The Reed-Solomon concatenation only adds acquisition time. Since the required BER performance of 10^{-3} is met and the MCD will not maintain lock below 7.0 dB, the added Reed-Solomon coding gain cannot be utilized to enhance threshold performance.

The total system modulation-demodulation loss for convolutionally encoded telemetry at 4.49 BPS is 4.0 dB. System loss for the uncoded 4.0 BPS is 2.4 dB. Noting that the carrier tracking margin is between 5.5 and 4.6 dB these losses appear reasonable.

A closing note should be made about the Block-IV SDA now in operation at DSS-14 and DSS-43. These advanced SDAs have third-order tracking loops which can significantly

reduce acquisition time. No calculations have been submitted at this time, the question has been left open for a potential future effort.

Reference

1. Rollins, W. W., "Confidence Level in Bit Error Rate Measurement," *Telecommunication and Data*, December 1977.

Abbreviations

BER	bit error rate
BPS	bits per second
CTA-21	Compatibility Test Area-21 at JPL
dB	decibel
DSN	Deep Space Network
E_b/N_0	energy per bit per spectral noise density
ISPM	International Solar Polar Mission
JPL	Jet Propulsion Laboratory
kHz	kilohertz
MCD	maximum likelihood detector
P_e	probability of error
PN	pseudo noise
RF	radio frequency
R-S	Reed-Solomon code
SCA	simulation conversion assembly
SDA	subcarrier demodulation assembly
SNR	signal to noise ratio
SPT	system performance test
SSA	symbol synchronizer assembly

Table 1. Bit error rate tabulated test data

Test no	Input E_b/N_0 , dB	Bit rate, BPS	Coding	Bit error rate, P_e	No of bits	System loss, dB	Carrier margin, dB
1	6.5	4.00	UC ^a	1.1×10^{-2}	52	2.4	4.6
	6.5	4.00	UC	1.1×10^{-2}	7	2.4	4.6
2	5.5	4.00	UC	2.2×10^{-2}	128	2.6	3.6
3	7.5	4.00	UC	5.0×10^{-3}	5	2.4	5.6
4	8.0	4.49	C ^b	1.2×10^{-4}	11	4.2	6.5
5	8.5	4.49	C	2.5×10^{-5}	94	4.2	7.0
6	7.0	4.49	C	1.0×10^{-3}	95	4.0	5.5

^aUC = uncoded

^bC = convolutional (7, 1/2, 3)

Table 2. Calculated threshold acquisition times for the RF and telemetry string

Subsystem	$E_b/N_0 = 7.0$ dB			$E_b/N_0 = 9.2$ dB		
	Uncoded	Coded		Uncoded	Coded	
RF (assumed)	00 10 00 ^a	00 10 00	—	00 05 00	00 05 00	—
SDA	00 36 40	01 33 20	—	00 20 50	00 46 40	—
SSA	00 02 28	00 02 12	—	00 02 28	00 02 12	—
MCD	—	00 03 34	—	—	00 03 34	—
Subtotal			01 49 06			00 57 26
R-S (1/2 frame)	—	00 20 50	—	—	00 20 50	—
R-S (4 frames)	—	—	00 42 00	—	—	00 42 00
Total	00 49 08 ^b	02 09 56	02 31 06	00 28 18	01 18 16	01 39 26

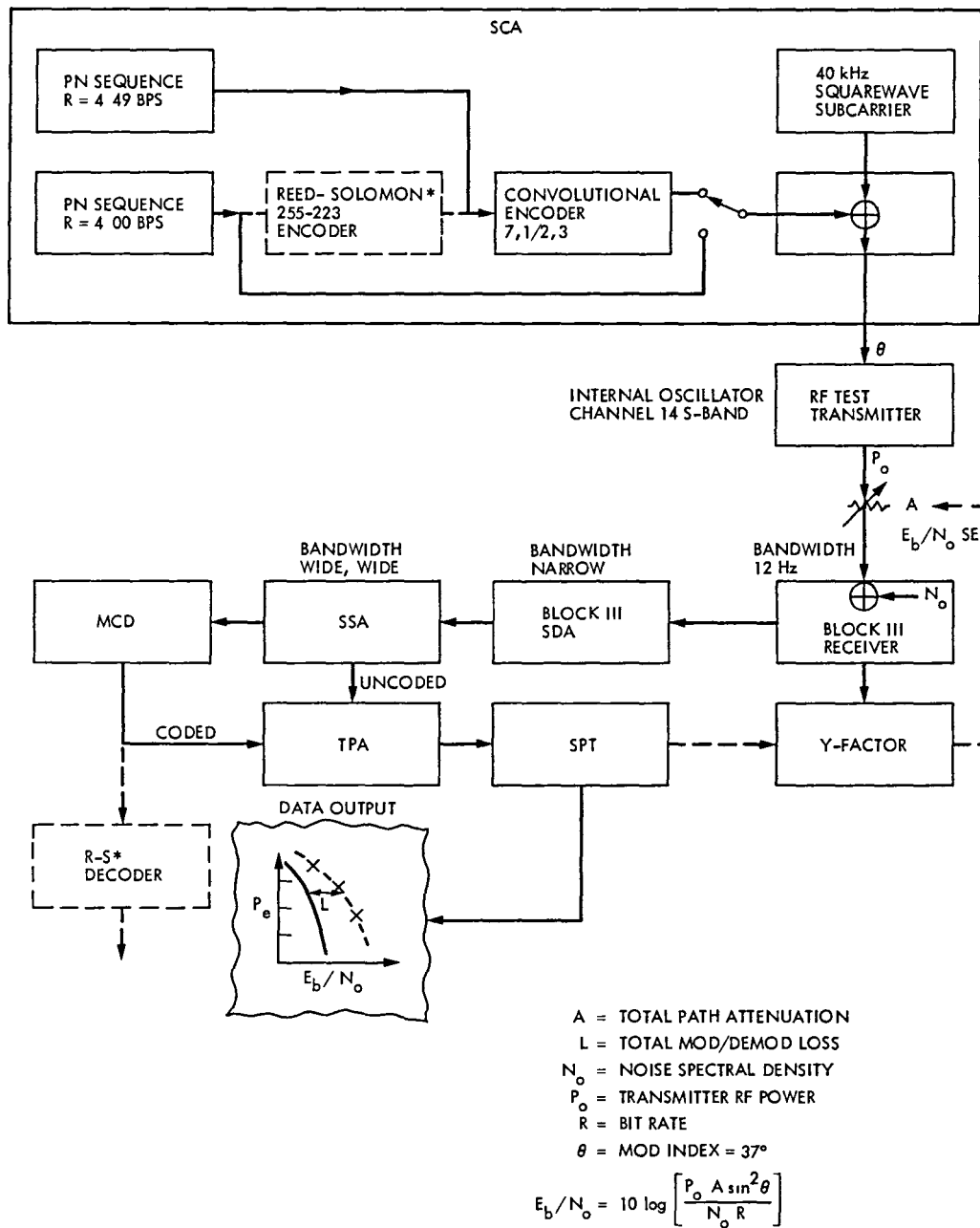
^aHours minutes seconds

^bBER performance is above specified threshold ($P_e > 10^{-3}$)

Table 3. Summary of acquisition time and BER performance for convolutionally coded and uncoded threshold E_b/N_0 input levels

Code conditions	$E_b/N_0 = 7.0$ dB		$E_b/N_0 = 9.2$ dB	
	Acquisition time	BER	Acquisition time	BER
Uncoded	00 49 08 ^a	8×10^{-3}	00 28 18	1×10^{-3}
Convolutional	01 49 06	1×10^{-3}	00 57 26	7×10^{-6}
Concatenated Reed-Solomon and convolutional	02 09 56 to 02 31 06	"Errorless"	01 18 26 to 01 39 26	"Errorless"

^aHours minutes seconds



*DUE TO PROGRAM RESTRICTION TO BIT RATES AT OR ABOVE 4 BPS, A BIT RATE OF 4.0 BPS WAS USED TO APPROXIMATE THE BIT RATE OF 3.93 BPS. CTA-21 HAS NO CAPABILITY FOR REED-SOLOMON

Fig. 1. Telemetry test configuration at CTA-21

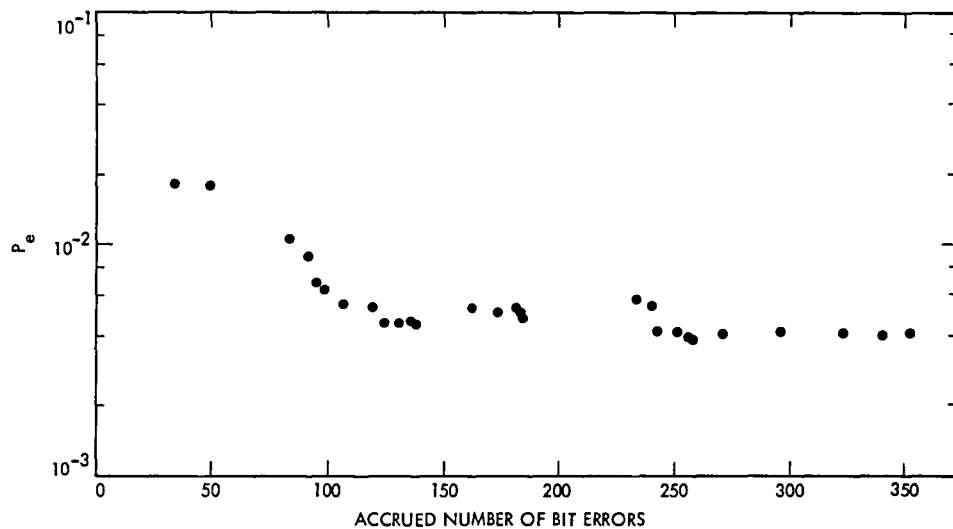


Fig. 2. Asymptotic behavior of accrued errors occurring in a high-speed block size of 960 bits, with input $E_b/N_0 = 6.5$ dB, $R = 4.49$ BPS, a convolutional code (7, 1/2, 3), and the carrier margin in $2B_L = 5.5$ dB

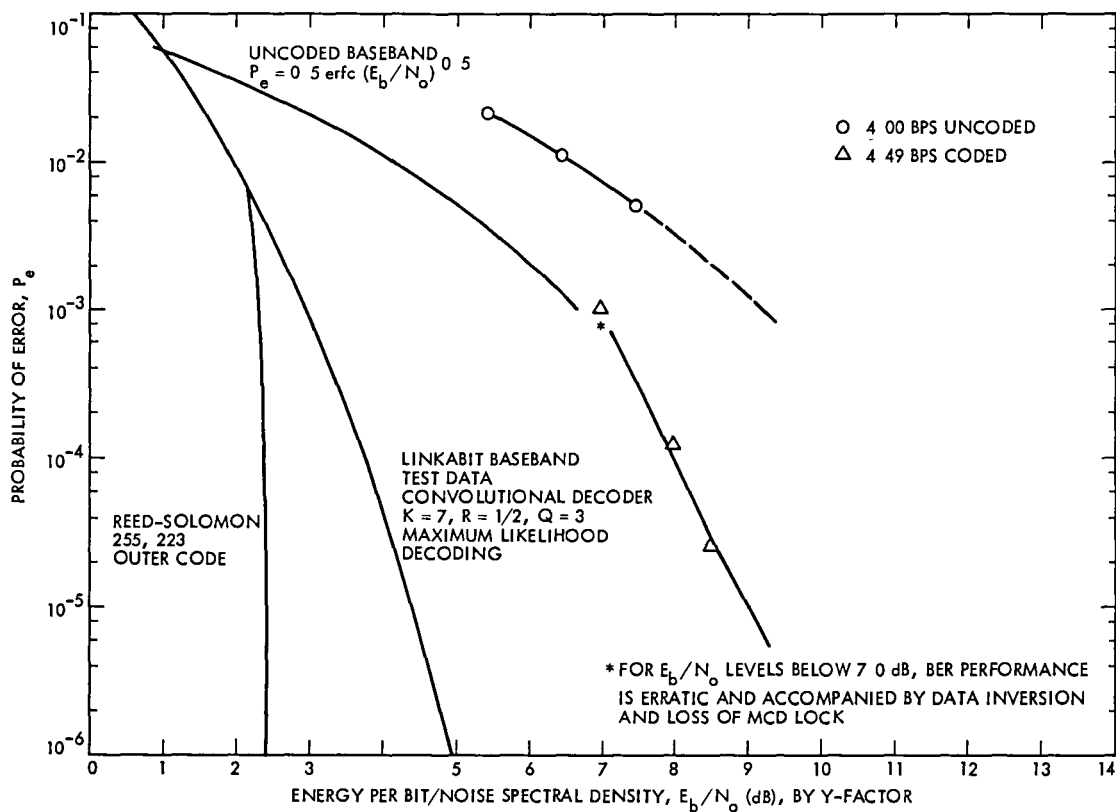


Fig. 3. NASA-ISPM emergency mode telemetry performance

Goldstone STDN 9-Meter Radiation Test

J R Blain

Deep Space Network Support Section

The GSTDN 9-meter tests were conducted from February through July 1981 to characterize the near-field radiation patterns of the S-band and fourth harmonic frequency emissions. The test configurations and results are presented, with graphs of the antenna patterns. The tests indicated that X-band leakage may be suppressed to levels of approximately -190 dBm/cm² at 200 meters

I. Introduction

The Mark IVA configuration of the Deep Space Network (DSN) will require that the antennas of each Deep Space Communications Complex (DSCC) be colocated close to the 64-meter antenna, except for Deep Space Station (DSS) 12 at Goldstone. This colocation of antennas within line-of-sight of each other and some as close as 200 meters from each other greatly increases the requirements for electromagnetic compatibility within the DSCC to insure continued mission support. Each of the antennas to be colocated represents a potential source of electromagnetic interference.

The DSN 34- and 64-meter antennas are presently colocated in Australia and Spain and have been engineered to suppress electromagnetic interference. However, the Mark IVA configuration also requires the colocation of a Spaceflight Tracking and Data Network (STDN) 9-meter antenna with transmit and receive capability for supporting near-earth missions. The 9-meter antenna has a diverse S-band operating frequency range, with a fourth harmonic that falls within the DSN X-band receiver operational frequency range. Further,

the 9-meter antenna tracks earth satellites at relatively fast tracking speeds, which substantially increases the opportunity for a transmitting antenna to radiate into the front of a receiving antenna. This made the 9-meter antenna the most likely antenna to cause electromagnetic interference (EMI). Existing data on the 9-meter antenna did not provide adequate information on its near-field radiation characteristics at fundamental and harmonic frequencies. The RF signature of this antenna was required to develop an antenna interference model. Therefore, actual near-field measurements of the STDN 9-meter's fundamental frequency and fourth harmonic radiations were required.

The STDN 9-meter facility normally operates in one of two modes: exciter bypass or power amplifier mode. In the exciter bypass mode, the output of the exciter bypasses the power amplifier (klystron) and is connected to the antenna through most of the same waveguide as the output of the power amplifier. Our initial concerns for these tests related only to the power amplifier mode, however later we had to include the exciter bypass mode in our test plan.

II. Test Criteria

Test criteria were as follows

- (1) Measure the maximum continuous wave (CW), radio frequency (RF) radiation from the 9-meter antenna, with the transmitter set for the frequency of 2105 MHz and the antenna radiating 10 kW, from a distance of less than 500 meters. Measurements were to be made at 2105 and 8420 MHz with the receiving antenna on boresight of the 9-meter antenna.
- (2) Measure the CW RF radiation levels of frequencies 2105 and 8420 MHz, as received from the 9-meter antenna, as a function of the antenna elevation in degrees relative to the boresight of the receive antenna. Transmit frequency 2105 MHz.
- (3) Measure the CW RF radiation levels of frequencies 2105 and 8420 MHz, as received from the 9-meter antenna, as a function of antenna azimuth in degrees relative to the boresight of the receive antenna. Transmit frequency 2105 MHz.

III. Test Configurations

A. 500-Meter Test Configuration

- (1) Test site 476 meters from the 9-meter antenna
- (2) Equipment configuration 1 (see Appendix).
- (3) Test system's signal measurable threshold ¹
 - (a) -134.5 dBm/cm² at frequency 2105 MHz
 - (b) -176.8 dBm/cm² at frequency 8420 MHz
- (4) Path loss ²
 - (a) S-band estimated to be 92.45 dB
 - (b) X-band estimated to be 104.49 dB

B. 200-Meter Test Configuration

- (1) Test site 200 meters from 9-meter antenna
- (2) Equipment configuration 2 (see Appendix).
- (3) Test system's signal measurable threshold.
 - (a) -172.0 dBm/cm² at frequency 8420 MHz

¹A preamplifier was locally fabricated to increase the sensitivity of the Hewlett-Packard, HP-8566A, that was used as the receiver

²All test locations are considered to be in the near field of the antenna radiation pattern. Therefore, the distance loss applies to signals in directions out of the main antenna beam

(b) S-band measurements were not required.

- (4) Path loss at X-band estimated to be 96.67 dB.

C. 70-Meter Test Configuration

- (1) Test site 70 meters from 9-meter antenna
- (2) Equipment configuration 3 (see Appendix).
- (3) Test system's signal measurable threshold was -172 dBm/cm² at frequency 8420 MHz
- (4) Path loss at X-band estimated to be 87.72 dB.

D. Equipment Leakage Test Configuration

- (1) Test site inside 9-meter transmitter equipment room.
- (2) Test system's signal measurable threshold
 - (a) -135 dBm at frequency 2105 MHz.
 - (b) -133 dBm at frequency 8420 MHz

IV. Test Results

A. 500-Meter Test Results

The S-band tests were performed and adequate information was obtained to predict the RF radiation pattern of the 9-meter antenna. See Figs. 1 and 2 for the radiation patterns. No further S-band testing was performed due to time restrictions.

The X-band tests were not completed due to time restrictions. However, the highest fourth harmonic power flux density levels were -122 dBm/cm² at 476 meters from the 9-meter antenna. Confirmation testing was conducted at the STDN 9-meter antenna facility in Spain with similar results. Investigations there concluded that the high level of radiation from the STDN antennas was due to a zero delay device³ (ZDD) that was bolted onto the side of the STDN antenna. When the device was removed, in Spain, from the 9-meter STDN antenna, the fourth harmonic signal level dropped below the detecting equipment threshold. The Goldstone fourth harmonic preliminary test results are depicted in Figs. 3 and 4.

B. 200-Meter Test Results

The primary test location was 200 meters from the Goldstone 9-meter STDN antenna. Fourth harmonic measurements

³The zero delay device is a small parabolic antenna, with a crystal detector, that is pointed at the subreflector of the 9-meter antenna. It is used for calibrating transmit and receive delay times to zero distance.

were conducted with the zero delay device (ZDD) installed and with the device removed. The difference in fourth harmonic signal level was dramatic, the maximum X-band signal level received with the ZDD installed was -110.5 dBm/cm^2 and maximum without ZDD was -152.2 dBm/cm^2 (see Fig 6). A total of four azimuth and four elevation tests were conducted with good repeatability. The test results are depicted in Figs. 5 and 6. The results of these tests indicated that, when the ZDD was removed, there was no discernible antenna pattern for fourth harmonic radiation.

After the last elevation test, the transmitter was terminated into a dummy load and a level of -159.5 dBm/cm^2 was received at the test site. Therefore we concluded that the fourth harmonic radiation signature measured when the ZDD was not installed was distorted by X-band leakage from within the 9-meter facility and/or reradiation. These distortions had to be reduced to measure the fourth harmonic antenna radiation signature. Therefore, additional tests were scheduled and the test site was moved to a location 70 meters from the 9-meter antenna. We wanted to find out whether the leakage could be eliminated by temporary fixes.

C. 70-Meter Test Results

A location approximately 70 meters from the 9-meter antenna was selected to enable continued illumination of the receive antenna by the entire 9-meter antenna facility, while minimizing adjacent reradiating surfaces. In conjunction with this test the spectrum analyzer was also used in the 9-meter transmitter facility to measure the exciter's output. Additional tests were made with the spectrum analyzer and a small X-band horn for detection of X-band radiation leaks in the transmitter system. See Table 1 for test data.

- (1) These tests validated the distance related model as the fourth harmonic signal levels initially measured at the 70-meter location were consistent with the previous measurements.
- (2) Almost equal levels of fourth harmonic signals were received at the 70-meter site, with transmitter terminated into the antenna or a dummy load. This verified that X-band leakage was distorting our ability to measure the X-band signature of the 9-meter antenna. Therefore, we conducted tests in the transmitter facility to isolate the leakage.
- (3) In the transmitter room, the main source of fourth harmonic generation was the exciter. The exciter had X-band components in its output and the exciter was radiating X-band RF signals from its power output stage.

- (a) The X-band components in the exciter's output varied 40 dB, while the S-band output remained within established parameters. Measured X-band levels varied from -98 to -58 dBm over the test period.
 - (b) The X-band radiation from the exciter cabinet was -77 dBm/cm^2 . The main radiator was the output stage which radiated -57 dBm/cm^2 , with exciter drawer open and output assembly exposed. Therefore, the exciter cabinet assembly only provided approximately 20 dB shielding.
 - (c) A notch filter (2100-2150 MHz, 50-dB attenuation) was used to determine the effect the exciter X-band output had on the X-band output of the klystron. These tests indicated that X-band input levels below -100 dBm had little effect on the klystron X-band output, while input levels of -100 to -70 dBm seemed to pass through the klystron increasing the X-band output by close to a one to one ratio.
- (4) The following was determined by leak detection testing.
 - (a) The transmission waveguide was leaking X-band signals at most flanges, and at the waveguide switch.
 - (b) The klystron was generating X-band signals.
 - (c) The highest detectable level of X-band radiation on the surface of the dish was -144.3 dBm/cm^2 , with 10 kW, S-band power output.

V. Temporary Fixes and the Differential Effects (see Table 1)

Trial fix	Differential effect	Test configuration
(1) Sealed a waveguide flange	Reduced X-band leakage 13 dB from -163.3 to -176.9 dBm/cm^2	6-26-81 70-m site, antenna at zenith, 10 kW output
(2) Sealed exciter cabinet	Reduced X-band leakage by 15 dB from -142.4 to -157.1 dBm/cm^2	7-2-81 70-m site, antenna at zenith, 10 kW output
(3) Sealed exciter cabinet and filtered exciter X-band components	Reduced X-band leakage by 25 dB from -146.6 to -174 dBm/cm^2	7-2-81 70-m site, TX antenna pointed at RX antenna, exciter bypass mode

VI. Conclusions

The tests characterized the S-band and fourth harmonic radiations of the 9-m antenna in its current operational configuration. The temporary fixes indicated that the X-band radiation levels may be suppressed to a satisfactory level.

A. Exciter Bypass Mode

By the use of temporary fixes we were able to reduce the level of fourth harmonic emission from the 9-meter antenna to a level below -172 dBm/cm^2 at 70 meters from the antenna. Also the X-band leakage was reduced to a level below -177 dBm/cm^2 .

The X-band output of the exciter varied more than 40 dB during these tests, while the S-band output was within tolerance.

B. 10-kW Power Amplifier Mode

The changes in the X-band components of the exciter's output caused variances in the level of fourth harmonic signals received at the test locations. The best case test results indicated a level of -168.5 dBm/cm^2 at antenna boresight and -176.9 dBm/cm^2 with antenna at zenith. Therefore, if the exciter X-band outputs can be controlled, the fourth harmonic power flux density will be below -170 dBm/cm^2 at 200 meters.

Acknowledgment

The author would like to acknowledge and thank Leslie G. Patterson for his support and guidance in conducting these tests. A special note of thanks to Gil Skipper and all the personnel of the GSTDN 9-meter facility for their cooperation and assistance. Technical guidance was provided by Dr. Ted Peng and Ray Caswell.

Table 1. 70-meter site test results data

Exciter bypass mode TX antenna pointed at RX antenna

Date, 1981	X-band PFD, dBm/cm ²	Test conditions
6-26	-163.4	WG flange "sealed"
6-30	-150.4	WG flange "sealed"
7-2	-174.0	WG flange and the exciter cabinet "sealed", exciter output filtered

Exciter bypass mode TX antenna pointed at zenith

6-26	-168	WG flange "sealed"
6-30	-150.8	WG flange "sealed"
7-2	-146.6	WG flange "sealed"
7-2	-175.8	WG flange and the exciter cabinet "sealed"
7-2	>-177	WG flange and the exciter cabinet "sealed," exciter output filtered

10-kW power amplifier mode TX antenna pointed at RX antenna

6-26	-168.5	WG flange "sealed"
6-30	-150.4	WG flange "sealed"
7-2	-154.5	WG flange and the exciter cabinet "sealed"

10-kW power amplifier mode TX antenna pointed at zenith

6-26	-163.3	Initial check
6-26	-176.9	WG flange "sealed"
6-30	-150.8	WG flange "sealed"
7-2	-142.4	WG flange "sealed"
7-2	-157.1	WG flange and the exciter cabinet "sealed"

Exciter RF power outputs

Exciter 1

Date, 1981	S-band, dBm	X-band, dBm	Remarks
6-17			Equipment inoperative
6-19	43	-74	Equipment repaired
6-30	43	-61	Equipment tuned
7-2	42	-58.1	No maintenance

Exciter 2

Date, 1981	S-band	X-band	Remarks
6-17	43	-71	Operating
6-19	43	-98.4	Equipment tuned
6-30	42	-83	Equipment tuned
7-2	42	-79	No maintenance

- NOTES**
1. Exciter output power was measured directly with spectrum analyzer using a band-pass filter.
 2. The remarks reflect actions taken by GSTDN personnel for their normal operation. No maintenance was performed during tests.

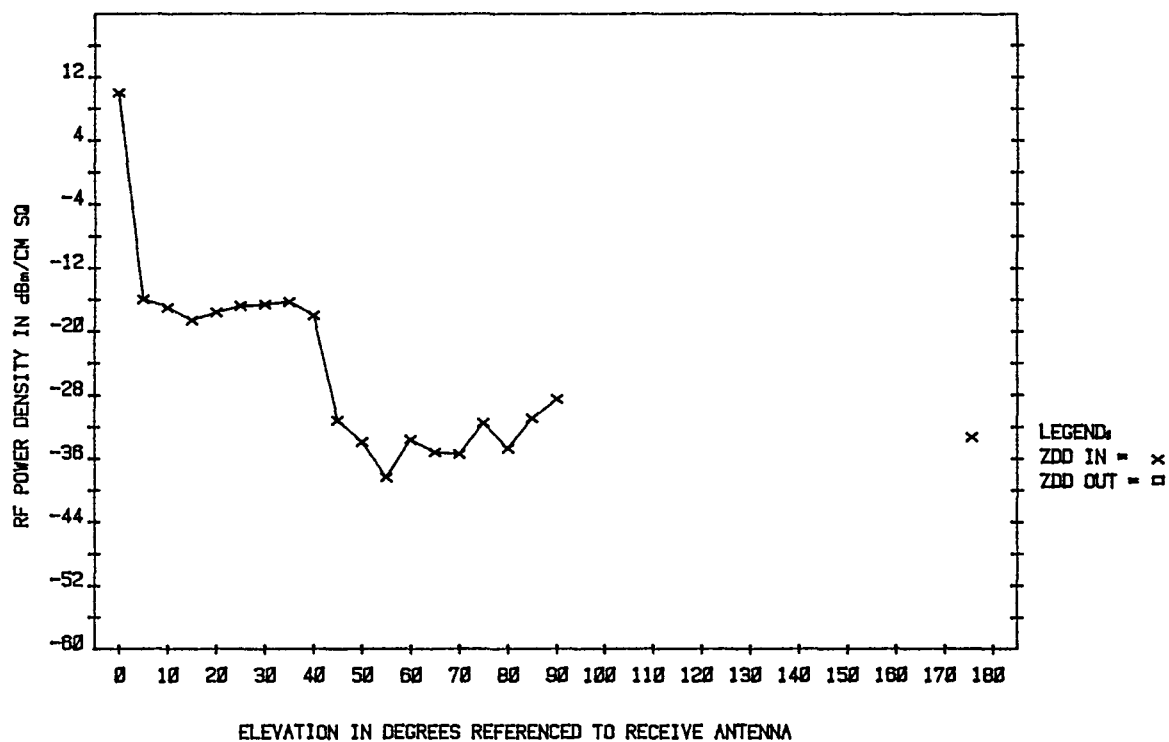


Fig. 1. Power flux density of S-band signal (2105 MHz) at 476 meters, elevation

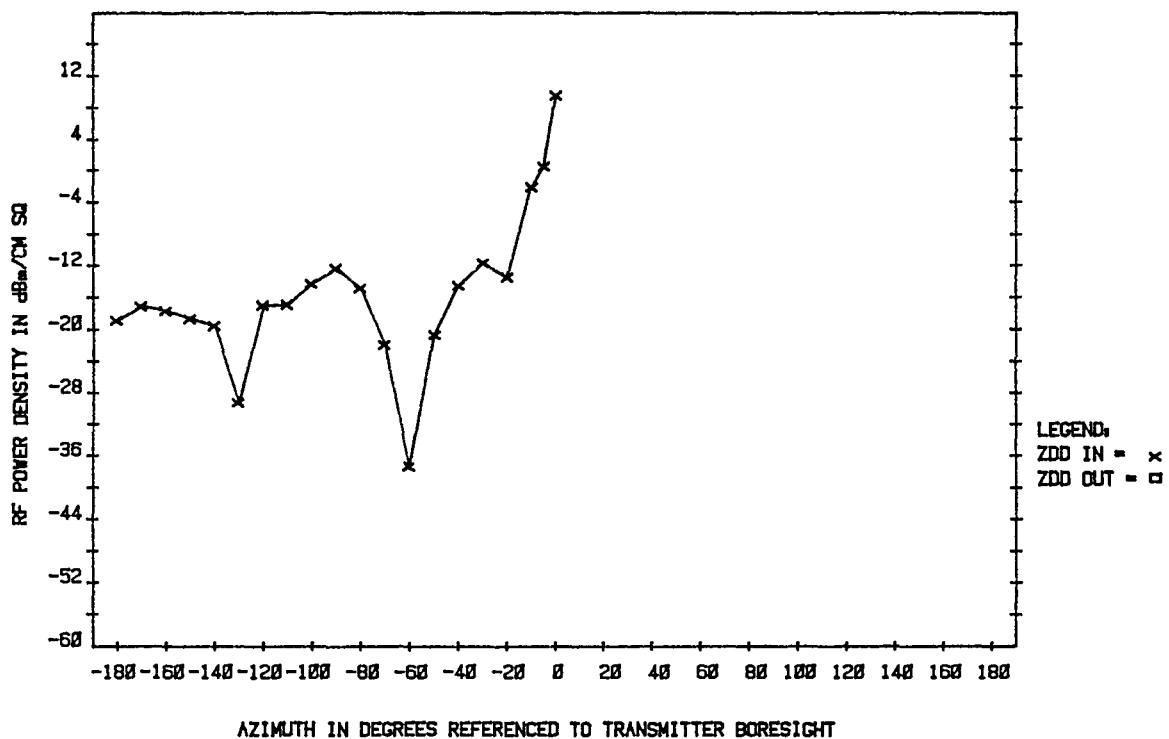


Fig. 2. Power flux density of S-band signal (2105 MHz) at 476 meters, azimuth

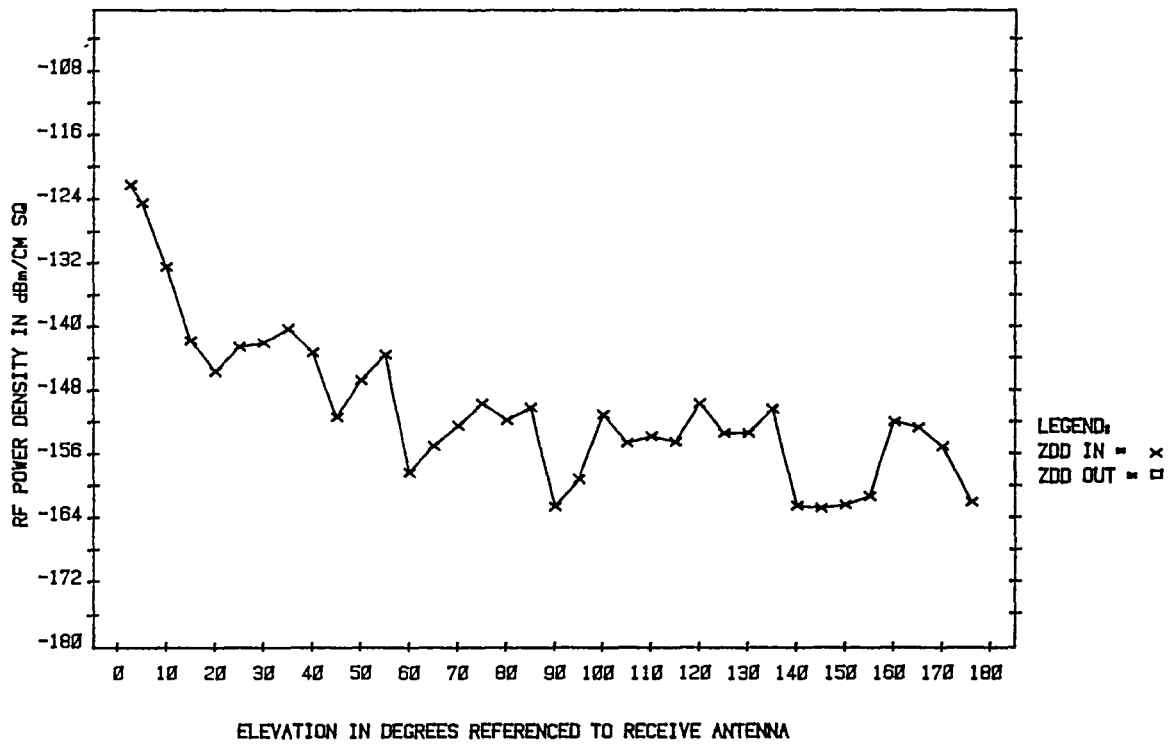


Fig. 3 Power flux density of fourth harmonic (8420 MHz) at 476 meters, elevation

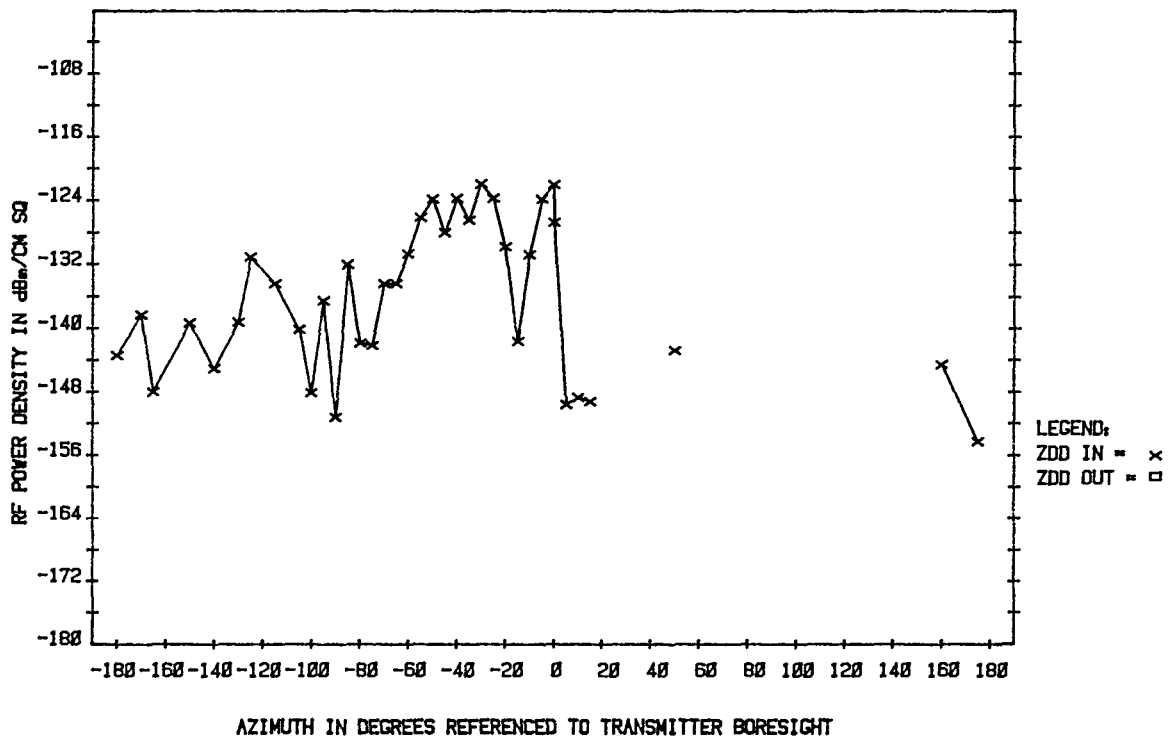


Fig. 4. Power flux density of fourth harmonic (8420 MHz) at 476 meters, azimuth

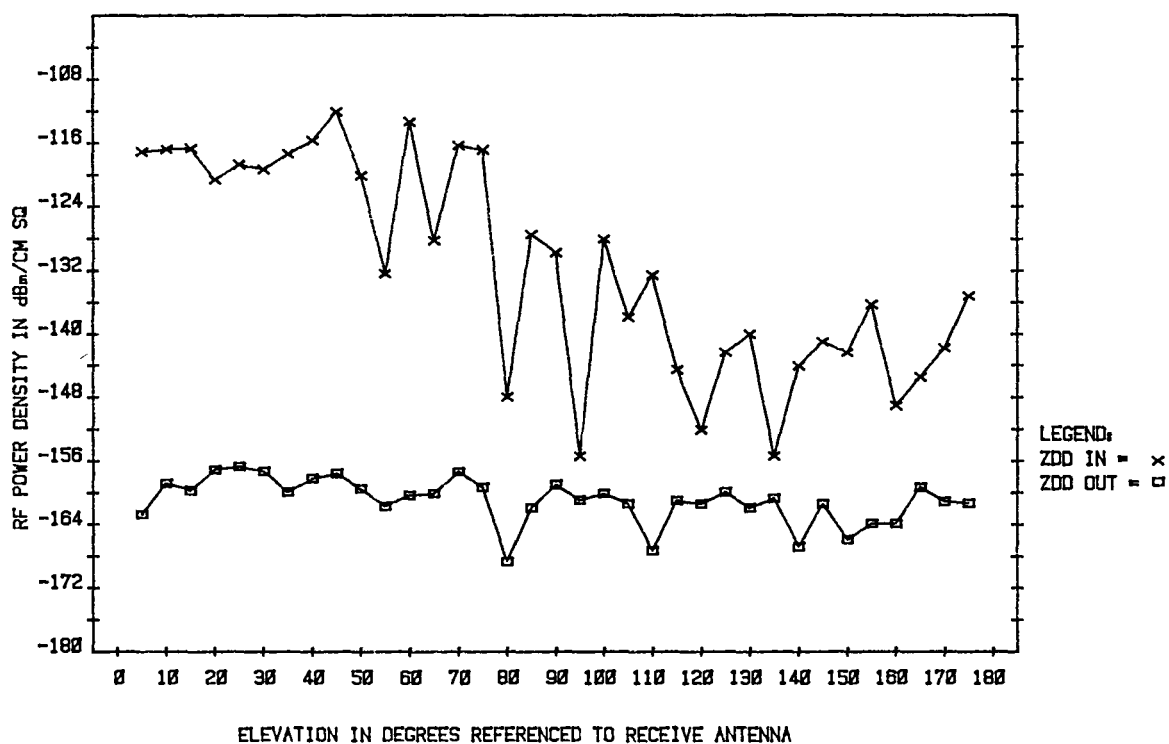


Fig. 5. Power flux density of fourth harmonic (8420 MHz) at 200 meters, elevation

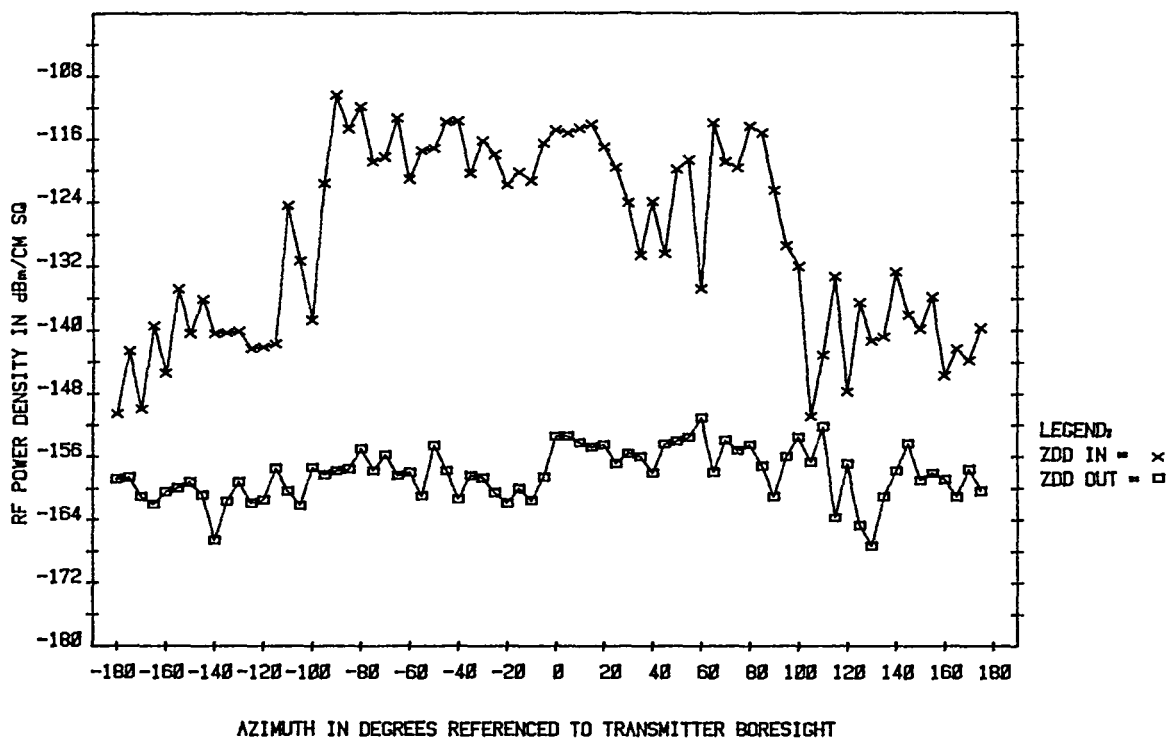


Fig 6 Power flux density of fourth harmonic (8420 MHz) at 200 meters, azimuth

Appendix

The S-band radiation detection equipment configuration is depicted in Fig. A-1. The S-band minimum measurable signal, with 5 dB signal-to-noise ratio was -134.5 dBm/cm². The X-band radiation detection equipment configuration is depicted in Fig. A-2.

To determine the minimum measurable signal of the X-band RF detection system employed for this project, we must first determine the temperature of the system. This RF detection system consists of

- (1) Antenna X-band, DSN standard gain, horn
- (2) Low noise front end Locally engineered X-band pre-amplifier.
- (3) Receiver Hewlett-Packard, HP-8566A, spectrum analyzer

System parameters are depicted in Fig. A-2. The loss L_3 was a variable due to different cable lengths required for specific measurement setups. The theoretical system temperature calculations are

System temperature

$$T_S = T_i + T_r$$

where

T_i = input temperature

T_r = receive system temperature

$$T_i = \frac{290}{L_1} \cdot 290 \left(1 - \frac{1}{L_1} \right) = 290 \text{ kelvin}$$

$$T_S = 290 + (F - 1) 290 + \frac{(F_2 - 1)_{290}}{G_1 \cdot L_2} + \frac{(F_3 - 1)_{290}}{G_1, G_2, L_2, L_3}$$

The cable length between the preamplifier and the spectrum analyzer resulted in the following values for L_3

Test Configuration 1 $L_3 = -5.6 \text{ dB}$

Test Configuration 2 $L_3 = -30 \text{ dB}$

Using the above formula and the parameters from Fig. A-2, we can derive the system temperature in kelvin. Once T_S is

derived we can determine the receive threshold by

$$\text{Noise threshold} = N_0 = KT$$

$$= -198.6 \text{ dB M/Hz} \cdot K + 10 \log(T_S)$$

and

$$\text{Receiver threshold} = N_0 BW = N_0 + \log(\text{RX bandwidth})$$

Therefore

Test Configuration	System Temperature	Receiver Threshold
1	473.8 K	-161.8 dBm
2	1,387.4 K	-157.2 dBm

To determine the minimum measurable input signal level, the RF detection system was set up in a certified screen room at the JPL Compatibility Test Area (CTA-21). An external, calibrated, X-band RF signal source (DSN X-band translator) was routed through precision attenuators and input to the screen room. By experimentation it was determined that, by employing 20 video averages, a signal-to-noise ratio of 5 dB could accurately measure the signal level.

As the RF detection system was to be used in the field, we assumed that minor antenna pointing errors may occur. Therefore we adjusted our antenna gain figure to 21 dB.

In view of the above, we concluded that, conservatively speaking, our thresholds including antenna gain were

Test Setup	Measurable Signal Threshold	Observable Signal Threshold
1	-176.8 dBm/cm ²	-181.8 dBm/cm ²
2	-172.0 dBm/cm ²	-177.0 dBm/cm ²

Thresholds in terms of field strength and power flux density are as follows

The field strength values may be determined by:

$$E = V_r + K$$

where

$$E = \text{dBm} = \text{dB referenced to } 1 \mu\text{V}$$

V_r = received signal in dBm = PFD

K = antenna factor in dB/m = $20 \log (F) - G - 29.8 \text{ dB}$

Thus $K = 27.70624183$, $F = 8420 \text{ MHz}$

(Note For 50 ohm input $0 \text{ dBm} = 107 \text{ dB}\mu\text{V}$)

The power flux density = $PFD = P_R - A_E$

where

$$A_E = G \text{ dB}_i + \frac{\lambda^2}{4\pi} (\text{dB}) = G \text{ dB}_i - 39.95 \text{ dB/m}^2$$

or

$$A_E = 21 - 40 = -19 \text{ dB/m}^2$$

PFD = threshold in dBm - (-19)

Therefore

(1) PFD at thresholds are

Test Setup	Measurable Signal Threshold	Observable Signal Threshold
1	-137.8 dBm/m ²	-141.8 dBm/m ²
2	-132.0 dBm/m ²	-137.0 dBm/m ²

(2) Field strengths at thresholds are

Test Setup	Measurable Signal Threshold	Observable Signal Threshold
1	-42.09 dB μ V	-47.09 dB μ V
2	-37.29 dB μ V	-42.29 dB μ V

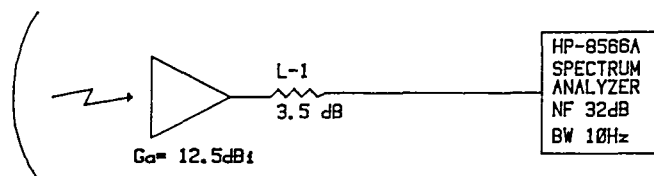


Fig. A-1 S-band equipment configuration

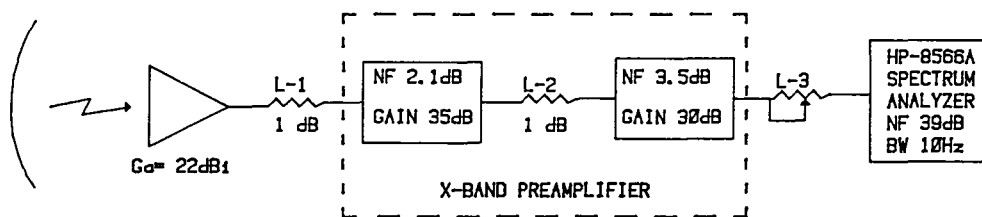


Fig A-2. X-band equipment configuration

A Solar Powered Water Purification System at Goldstone (DSS 13)

F. J. Menninger and R. J. Elder
DSS Engineering Section

The present energy intensive electrically powered still at the Venus Deep Space Station (DSS 13) has reached its end-of-life. The electric still has been producing the pure water required for cooling the 400-kW klystron tube of the Venus 26-m antenna at Goldstone, California. A new passive solar still, unique to the Goldstone complex, has been designed to replace the electric still. This new system will deliver 100% of the pure water requirement for cooling the klystron tube and will consume an insignificant amount of electrical energy. Construction will be completed by the end of 1981, followed by a one-year test and evaluation period to gather data on performance characteristics.

I. Introduction

At the Venus station (DSS 13) of the Deep Space Network (DSN) Tracking Complex at Goldstone, California, distilled water of a high electrical resistivity quality is required as a cooling medium in the microwave tube generator (klystron tube). An electrically powered distillation system to provide this pure water has been in service for over 15 years and as a result has reached its end-of-life. Although the electric still is currently in operation, the dramatic increases in energy and maintenance costs have led to the investigation of alternative configurations and sources of supply. Examples of alternate systems studied for technical feasibility and cost effectiveness include delivery from an independent commercial supplier of pure water to the geographically remote tracking facility, trucking in pure water from Los Angeles using a truck available from nearby Yermo, installation of a new electric still; a reverse osmosis system; and construction of a solar still.

After careful examination of the reliability and life cycle costs of the alternate approaches, a passive solar powered still

was found to be the most viable solution due to the high quality of water produced, low installation, maintenance and operating costs, and minimal energy consumption.

Solar water purification units have been utilized for many years in arid areas as an efficient and inexpensive method for the production of drinking water from salty or brackish sources. The basic principles by which a solar still operates are straightforward. The sun's rays enter through a sloping cover glass, warming the water and producing a vapor which condenses on the inner surface of the sloping cover. The water droplets coalesce and flow down the slope into a discharge trough as distilled product water (Fig. 1).

II. Design Requirements

To satisfy the requirements of quantity and quality of the pure water produced, the new solar still system has been designed to deliver distilled pure water with a resistivity of at least 50 k Ω per cm at the underground storage tank. This

system will be composed of eight 1 85-m² (20-ft²) solar still panels producing a total annual yield in excess of 22.7 m³ (6000 gal) of pure water. The system will also incorporate a time clock and a solenoid valve for totally automatic operation of the flush and fill functions. The system is gravity feed, utilizing a temporary holding tank to store pure water for release to the underground storage tank.

III. System Configuration and Operation

Each solar still panel is a molded fiberglass tray with glass glazing sealed with silicone rubber (RTV) as illustrated in Fig. 1. The bottom of the panel is insulated with rigid foam and reinforced with marine plywood. The base of the wetted tray is colored black to help accelerate the evaporation, and the balance is colored white. The eight solar still panels, oriented to present a south facing slope as shown in Fig. 2, are attached to unistrut and held down with steel clamps that are mounted to a welded steel frame. All piping interconnections are polyvinyl chloride (PVC) and stainless steel. The complete system is elevated 3.35 m (11 ft) above ground level to facilitate the gravity feed, eliminate shadowing, and place panels above excessive traffic flow.

Local well water (with a high mineral content) is supplied to the solar still system by gravity as shown in Fig. 3 from the existing elevated raw water storage tank. The storage tank outlet is 91 m (300 ft) above the inlet to the solar stills. Solar radiation produces distilled pure water in the solar panels which is then gravity fed to a 322-liter (85-gal) temporary holding tank for a period of two to four days. This water is then released to the existing 5.7-m³ (1500-gal) underground pure water storage tank. An automatic once-a-day cycle puts the still in a flush mode to remove accumulated mineral concentrations and provides fresh fill water for the next day's pure water production. Previous studies and operation indicate

that the efficiency of operation does not decrease if the pan is contaminated or the top glass becomes dirty. This eliminates any type of routine maintenance or cleaning.

IV. System Environment

The Goldstone NASA tracking station in the Mojave desert with an elevation of 914 m (3000 ft) has abundant solar radiation, clear sky and clean air, which provides for an ideal solar application. There are 2591 heating degree days per year and an average annual horizontal insolation of 5.81 kWh/m² day (1843 Bth/ft² day).

V. Costs

The solar still is a totally passive system that does not use any significant electrical power. The occasional use of electrical control power for a supply line solenoid, timer, and indicator light is the only power required.

The 1981 annual maintenance and operation (M&O) costs for the existing electric still, including the electrical energy consumption, are \$5.1K. Based on a 1981 M&O cost of \$20/1000 gallons for a solar still, the annual savings over operation of the electric still would amount to \$4.9K. With a 10% per year escalation factor in labor and electrical costs, the payback period for the \$24K construction and installation costs of the solar still will be 3.5 years. The rapid payback period with improved reliability and reduced M&O costs were driving points to size the system for 100% solar.

The project will be operational in October 1981. Twelve months of testing will follow to develop data pertaining to weather conditions relative to pure water production quantity, temperature, purity, etc.

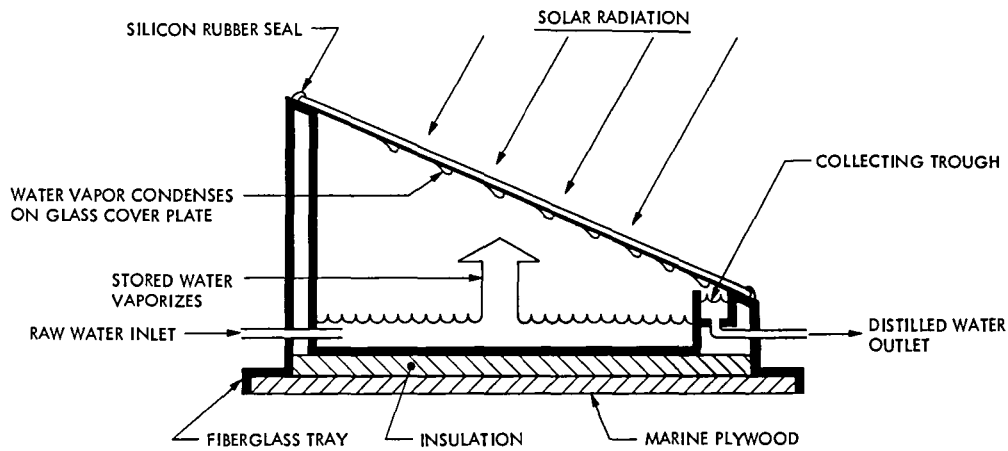


Fig 1. Solar still panel function and construction

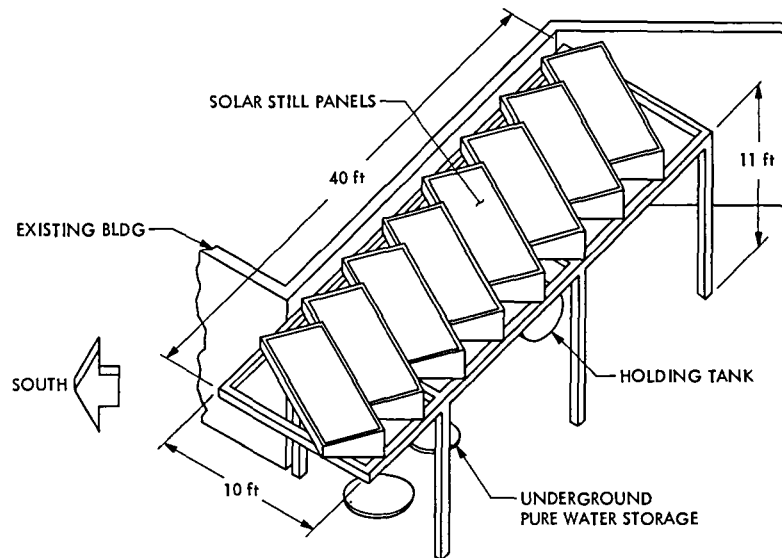


Fig 2 Solar still system installation

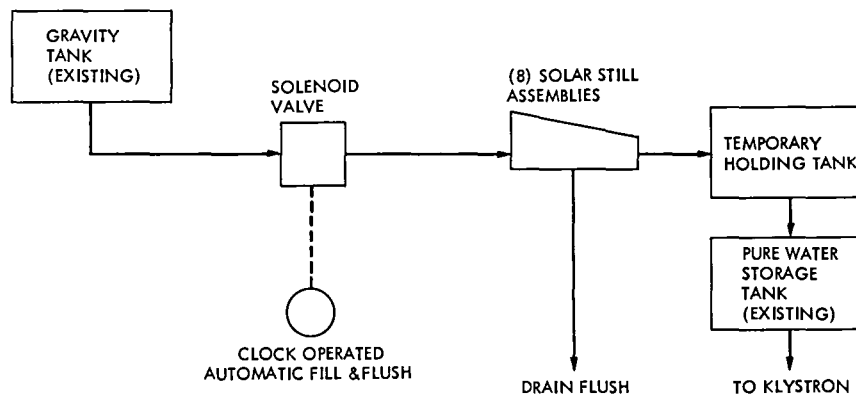


Fig 3 Solar still system flow diagram

Energy Consumption for the Echo Station (DSS 12)

C N Guiar and D Schonfeld
DSN Engineering Section

This article presents an analysis of the energy consumption for the Echo Station (DSS 12) HVAC equipment is shown to be the largest consumer of electrical energy. The energy consumption for the buildings at this station is itemized and compared to the electrical meter data. The figures for energy consumption as predicted by the ECP computer program are found to agree well with the meter data.

I. Introduction

At the end of 1973, the Goldstone Deep Space Communication Complex (GDSCC) initiated an energy management program, following NASA guidelines, towards energy consumption reduction and self-sufficiency. The program objectives, formally started in 1975, were to reduce the consumption of purchased energy by 50% by the end of 1985, using the consumption level in 1973 as a base. In 1977, a similar program was formally started for the overseas Deep Space Stations.

A reliable model for the simulation of energy consumption in buildings is necessary in order to implement this reduction of purchased energy. A computerized model for energy consumption in buildings is developed for this purpose, the ECP Program (described in Refs 1, 2, and 3), using the description of building construction, weather data, and mechanical and electrical systems. The ECP calculation of energy requirements for heating or cooling in any building involves three major consecutive steps. First, the heat loss or heat gain to the space that is heated or cooled is computed. Second, the

heating or cooling loads imposed on the heating or cooling coils inside the air-handlers are determined. Third, the energy input to all of the primary equipment or components constituting the air-conditioning system, such as compressors, heat pumps, boilers, and engines is calculated. Given these calculations, ECP can then provide an economic analysis to evaluate the payback period of money spent on a specific modification.

Initial verification of the computer model is done by comparing the results of the computer simulation to the measured meter data. The verification yields data pertaining to the two types of energy consumption: electrical and thermal, as well as daily, monthly, and yearly total building or site consumption.

The GDSCC is divided into a number of sites. In this article we present the verification analysis for the Deep Space Station Echo (DSS 12). This is the largest site and includes a 26-m antenna and 22 support, storage, and control buildings as listed in Table 1.

II. Energy Consumption

The ECP Program output provides a detailed simulation of the total energy consumption, dividing the energy consumers into the following groups: HVAC equipment, accessories, electronics, and lights. The HVAC equipment group includes vapor compression refrigeration units, absorption chillers, and heating units such as boilers and heat pumps. The accessories include the equipment that is necessary for building operation but does not affect heating and cooling load calculations. The accessories are subdivided into thermal and electrical. Included in the electric accessories are air-handler fans, condenser fans (air-cooled), condenser pumps (water-cooled), and boiler pumps. Thermal accessories equipment includes fuel consuming devices that are not located within any air-conditioned zones such as domestic hot-water boilers. The electrical and electronic equipment includes computers, electronic racks, and electronic equipment not related to air-conditioning and/or accessories. The lighting equipment is classified into incandescent and fluorescent types. Table 2 summarizes the simulated energy consumption for the entire Echo Station on a monthly basis.

A. Electrical Consumption Analysis

The distribution of energy between HVAC equipment, accessories, electronics, and lights is illustrated in Fig. 1. From Fig. 1, it can be seen that the largest consumers of electrical energy are the HVAC equipment (29.1%), electrical and electronics equipment (29.0%), and accessories (27.6%). The smallest consumers are the lights (10.1%) and machinery (4.3%). Figure 2 gives a graphical representation of the electric consumption for the entire Echo Station on a monthly basis. The pattern of the consumption shown in this figure is due mainly to the heating and cooling loads of the station. These loads are shown on a monthly basis in Figs. 3 and 4. The data on which these two figures are based is given in Tables 3 and 4.

In Table 5 the yearly electric consumption for each building in the station is summarized. Six buildings, G-29, G-36, G-39, G-40, G-43, and G-44, are not included in Table 5 because they are extremely small consumers of energy. For example, building G-29 has only two light bulbs and G-43 has been shut down. Figures 5, 6, 7, and 8 indicate which buildings are the largest consumers in each electrical category. G-26 is the largest consumer for electric accessories, electronics, and lights. G-38 is the largest consumer for electric equipment. Other large consumers in various categories are buildings G-33, G-23, and G-34/35. The relative distribution of building electrical consumption at the Echo Station is shown in Fig. 9.

B. Thermal Consumption Analysis

Table 6 shows the thermal consumption for the buildings forming the Echo Station. Thermal energy is obtained from the burning of LPG. As can be seen from Table 6 and Fig. 10, the largest thermal accessories consumer is building G-23. The largest equipment consumers are buildings G-33 and G-38, constituting over one-half the total consumption.

C. Verification of Results

A number of watt-hour meters monitor the electrical consumption at the Echo Station. A diagram of the metering system is shown in Fig. 11. There are two types of meters on the site: those for Southern California Edison (SCE) generated power and those for station generated power. The latter are meters 04 and 61; all others are SCE meters. Meters 04 and 60 are connected to the major tracking buildings. Meter 59 is connected to the smaller "housekeeping" buildings. All meters are read once a month. The meter readings are averaged for the years 1977 to 1979. This averaging process is necessary because of variations in tracking hours and, therefore, consumption. These readings are compared with ECP values and the results are given in Table 5 where individual and site comparisons between meter data and ECP values are made. Note that the largest difference between the two is about 10%, and for the complete site the discrepancy is less than 5%.

One meter monitors LPG consumption for the Echo Station (meter 19, shown on Fig. 11). Comparing the ECP thermal consumption of these buildings with the meter readings, averaged for the years 1977 to 1979, we observe that they agree within 6.64%.

III. Summary

It is important to realize that the GDSCC is a dynamic system with programmatic and nonprogrammatic equipment constantly being added, improved, and eliminated. Changes in building functions occur periodically; therefore, the present verification report represents only a basis for future analysis of the energy consumption of the Echo Station.

Comparison of the simulated consumption values with actual measured values shows ECP to be an accurate program when applied to energy consumption of the Echo Station. The program has identified the largest electrical consumers to be G-26 and G-38 and the largest thermal consumers to be G-23, G-33, and G-38. The verification analysis for the other Deep Space Stations at Goldstone will be presented in another TDA progress report.

References

1. Stoller, F W , et al , “Energy Consumption Program,” in *The Deep Space Network Progress Report 42-45*, pp 288-293, Jet Propulsion Laboratory, Pasadena, Calif , 1978
2. Lansing, F L , et al , “The Updated Algorithm of the Energy Consumption Program (ECP),” *The Deep Space Network Progress Report 42-49*, pp 107-115, Jet Propulsion Laboratory, Pasadena, Calif , 1979
3. Plankey, B , “Computer Simulated Building Energy Consumption for Verification of Energy Conservation Measures in Network Facilities,” *TDA Progress Report 42-62*, Jet Propulsion Laboratory, Pasadena, Calif., April 1981, pp 142-146

Table 1. Echo site buildings

Building No	Building description
G-21	Administration and Cafeteria
G-22	Pump House
G-23	Dorm
G-24	New Power Plant
G-25	Transportation and Shop
G-26	Control Building
G-27	Generator Building
G-28	Machine Shop
G-29	Storage Building
G-30	HVAC
G-33	Engineering and Communications
G-34	Hydromec
G-35	26-m Antennae
G-36	Collimation Tower
G-37	Seismic Lab
G-38	Lab and Maintenance
G-39	Paint Shop
G-40	Flammable Storage Building
G-41	Shipping, Receiving, and Warehouse
G-42	Maintenance Office
G-43	Security
G-44	Drum Storage Building
G-45	Carpenter Shop
G-46	Antennae Repair
G-47	Supply Warehouse

Table 2. Simulated energy consumption for Echo site using ECP Program^a

Month	Accessories		Lights		Electrical and electronic equipment, kWhe	Machinery, kWhe	HVAC Equipment	
	Thermal, kWht	Electric, kWhe	Incandescent, kWhe	Fluorescent, kWhe			Thermal, kWht	Electric, kWhe
January	70	99,291	9,334	28,757	109,852	12,969	169,206	57,937
February	66	90,351	8,450	26,346	99,512	11,854	124,696	69,690
March	70	98,166	9,334	28,757	109,852	12,969	128,868	81,887
April	70	95,715	9,059	28,100	106,529	12,644	113,722	91,888
May	70	100,797	9,334	28,757	109,852	12,969	114,302	103,457
June	70	109,899	9,059	28,100	106,529	12,644	82,557	142,108
July	70	115,513	9,284	28,626	109,632	12,963	86,602	152,822
August	70	116,563	9,383	28,887	109,971	12,750	88,230	150,510
September	70	107,988	8,960	27,839	106,199	12,632	88,375	134,940
October	70	101,655	9,334	28,757	109,852	12,969	103,494	116,227
November	70	99,044	9,059	23,100	106,529	12,644	111,181	71,620
December	70	102,268	9,087	28,103	108,992	12,939	165,769	53,594
Year Total	836	1,237,250	109,677	339,129	1,293,301	152,870	1,411,390	1,226,680

^aNumbers rounded to the nearest integer

Table 3 Heating load in MWh(e) for the Echo site

Building No /Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Building total
G-21	13.4	10.0	9.8	9.0	13.0	9.7	8.4	8.7	10.4	7.4	10.7	14.1	124.6
G-23	8.5	6.5	6.4	5.5	5.4	4.0	3.2	3.4	4.5	4.6	7.0	8.8	67.8
G-24	1.3	0.3	0.2	—	—	—	—	—	—	—	0.2	1.6	3.5
G-25	0.7	0.2	0.1	— ^a	0.2	— ^a	—	— ^a	— ^a	— ^a	0.2	0.9	2.4
G-26	22.6	16.0	16.5	13.9	9.9	5.4	4.1	4.3	6.8	12.4	17.6	23.8	153.3
G-27	4.9	2.8	2.8	2.2	1.6	1.0	0.5	0.5	1.2	1.5	3.2	5.3	27.6
G-28	0.4	0.2	0.2	0.1	0.3	0.1	—	—	0.1	—	0.2	0.4	2.1
G-30	0.1	—	—	—	—	—	—	—	—	—	—	0.1	0.1
G-33	36.9	28.7	30.4	27.3	29.2	22.8	20.8	21.2	24.2	25.3	31.0	38.1	335.8
G-34/35	0.1	—	—	—	—	—	—	—	—	—	—	0.2	0.2
G-37	1.5	0.6	0.5	0.3	0.1	—	—	—	—	—	0.6	1.7	5.3
G-38	55.4	47.4	51.9	50.5	56.1	44.9	53.9	54.2	45.9	56.1	50.6	55.2	622.3
G-41	0.4	0.1	— ^a	—	0.2	—	—	—	— ^a	—	0.1	0.5	1.3
G-42	0.1	— ^a	— ^a	—	— ^a	—	—	—	—	—	— ^a	0.1	0.3
G-45	0.6	0.5	0.4	0.3	0.2	—	—	—	— ^a	0.1	0.5	0.6	3.1
G-46	0.2	— ^a	—	—	— ^a	—	—	—	—	—	— ^a	0.3	0.6
G-47	0.5	0.3	0.2	0.1	0.2	—	— ^a	—	— ^a	— ^a	0.4	0.7	2.6
Monthly Total	171.1	103.6	119.3	109.3	116.5	88.0	91.0	92.3	93.2	107.6	122.4	152.1	

^aA value less than 0.1 MWh

Table 4 Cooling load in MWh(e) for the Echo site

Building No /Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Building total
G-21	6.4	8.4	9.5	10.5	11.7	15.3	17.0	16.6	14.4	12.2	8.6	5.8	136.3
G-23	0.8	1.5	1.8	2.3	0.9	2.8	3.9	3.6	2.0	2.8	1.5	0.6	24.5
G-24	15.8	14.6	16.2	17.9	30.3	33.2	36.5	36.2	32.0	19.8	15.6	15.7	283.9
G-25	— ^a	0.5	0.6	0.7	0.8	1.2	1.2	1.2	1.0	0.8	0.5	— ^a	8.6
G-26	76.8	72.7	81.4	80.5	87.5	89.7	95.4	94.9	88.1	85.3	77.5	76.0	1006.0
G-27	—	0.7	1.0	1.2	1.8	2.7	3.3	3.3	2.3	1.4	0.6	—	18.3
G-28	2.3	2.4	2.8	2.8	2.8	3.1	3.4	3.4	3.0	3.1	2.5	2.2	33.9
G-30	— ^a	0.3	0.4	0.5	0.6	0.9	1.1	1.0	0.8	0.6	0.3	—	6.6
G-33	36.2	37.5	43.3	44.4	48.0	54.5	60.6	60.5	51.4	49.4	39.9	34.1	559.7
G-34/35	2.7	4.0	5.1	7.4	15.3	24.6	30.4	29.9	22.2	10.8	3.9	2.5	158.9
G-37	—	0.2	0.3	0.5	0.8	2.0	2.7	2.6	1.6	0.9	0.1	—	11.8
G-38	52.3	53.3	60.0	62.7	76.7	75.4	91.1	90.6	73.5	74.2	55.9	50.1	815.6
G-41	—	1.3	1.9	2.8	1.3	3.7	4.6	4.3	2.7	3.5	1.0	—	27.1
G-42	0.1	0.2	0.2	0.3	0.2	0.4	0.5	0.5	0.3	0.4	0.2	— ^a	3.3
G-45	—	—	—	—	— ^a	0.5	0.7	0.7	0.4	0.1	—	—	2.4
G-46	—	0.1	0.3	0.4	0.3	0.6	0.8	0.7	0.5	0.4	0.1	—	4.2
G-47	—	—	—	0.2	0.2	0.8	1.0	1.0	0.6	0.5	— ^a	—	4.3
Monthly Total	193.5	197.7	224.9	235.2	279.1	311.2	354.3	351.0	296.9	266.3	208.2	187.1	

^aA value less than 0.1 MWh

Table 5 Total yearly electric consumption for Echo site buildings

Building no	Electrical consumption, kWh
G-21	224,673
G-23	30,860
G-24	628,664
G-25	48,388
G-26	1,109,720
G-27	38,870
G-28	93,899
G-30	8,694
G-33	705,700
G-34	473,659
G-37	15,571
G-38	941,875
G-41	53,453
G-42	27,917
G-45	23,528
G-46	11,363
G-47	30,796

Table 6. Thermal consumption for buildings at Echo site

Building	Accessories		HVAC equipment	
	kWh	Percent of site total	kWh	Percent of site total
G-21	836	100	155,758	11.0
G-22			15,245	1.1
G-23			84,879	6.0
G-24				
G-25			3,805	0.3
G-26			195,454	13.8
G-27				
G-28			2,807	0.2
G-30			350	0.025
G-33			419,786	29.7
G-34/35				
G-37				
G-38			527,750	37.4
G-41			1,594	0.1
G-42				
G-45			3,961	0.3
G-46				
G-47				
Total	836	100	1,411,390	100.0

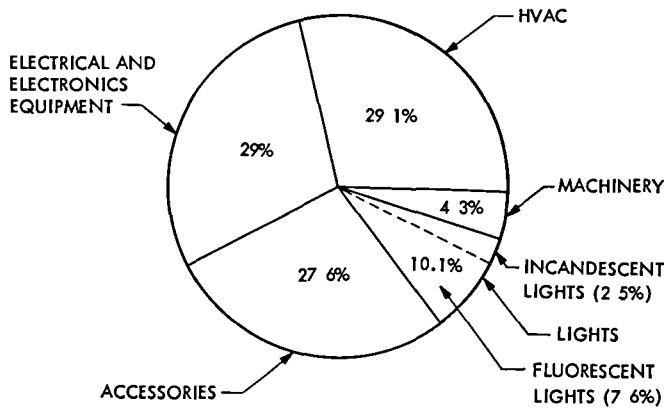


Fig 1 Itemization of annual electrical consumption at Echo Station

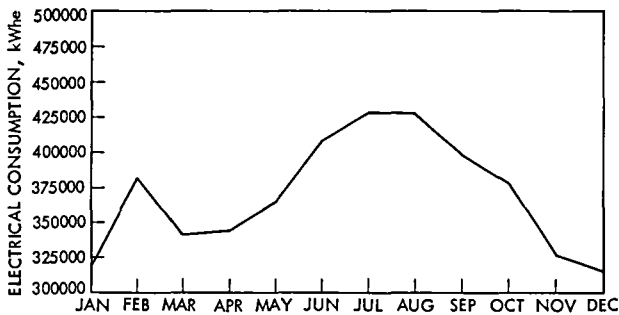


Fig 2. Electric consumption on a monthly basis

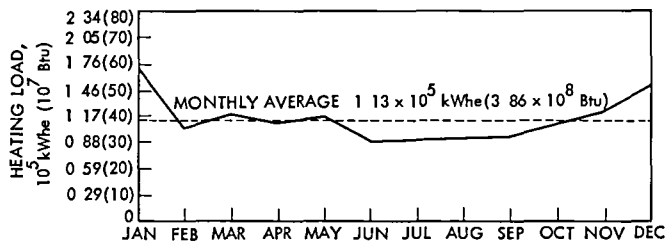


Fig. 3 Monthly variations in the heating load of Echo Station

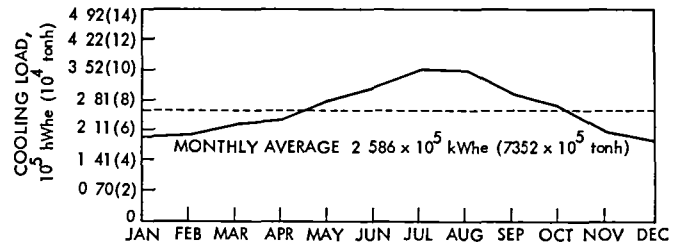


Fig. 4. Monthly variations in the cooling load of Echo Station

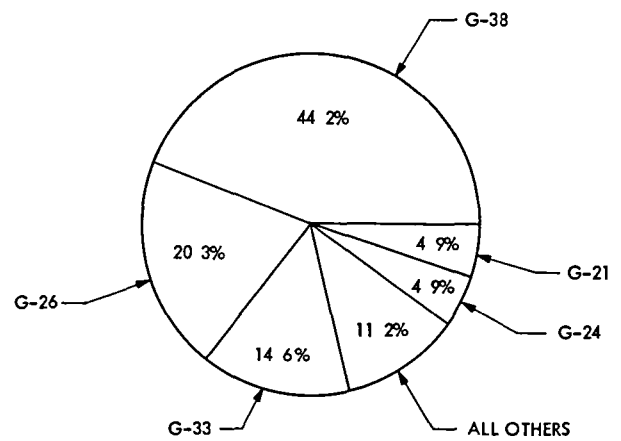


Fig. 5. HVAC electrical consumption for buildings at Echo Station

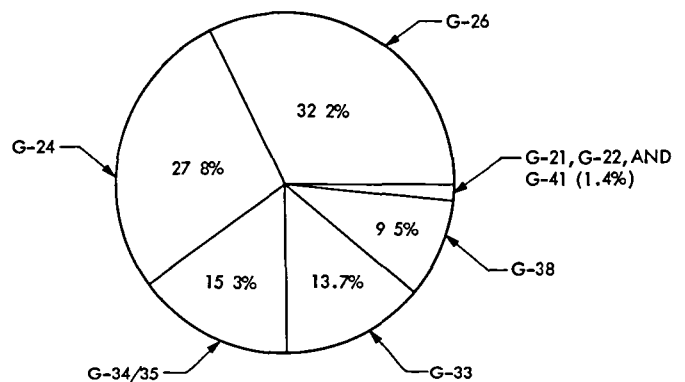


Fig. 6. Electrical and electronics equipment consumption for buildings at Echo Station

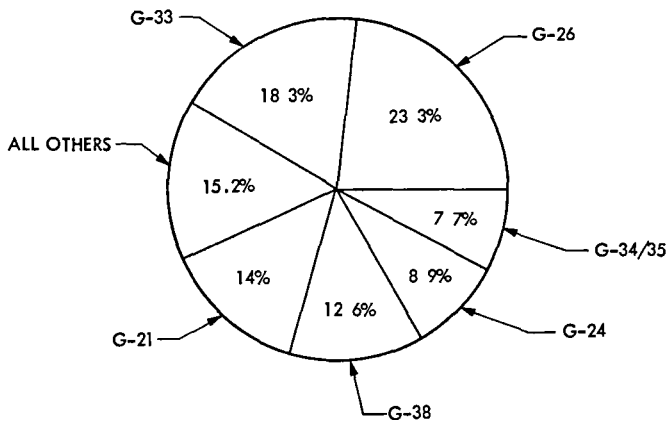


Fig. 7. Electric accessories consumption for buildings at Echo Station

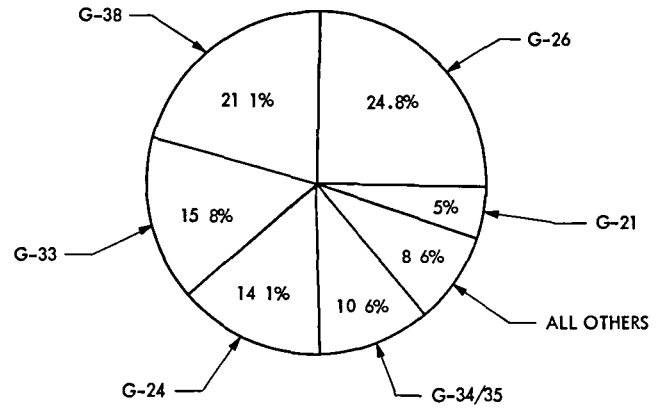


Fig 9 Total electrical consumption for buildings at Echo Station

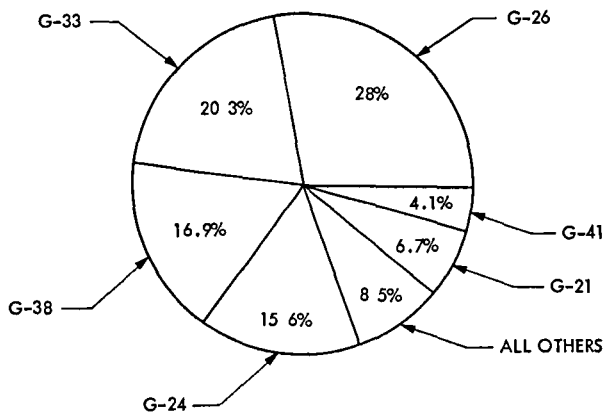


Fig 8. Light consumption for buildings at Echo Station

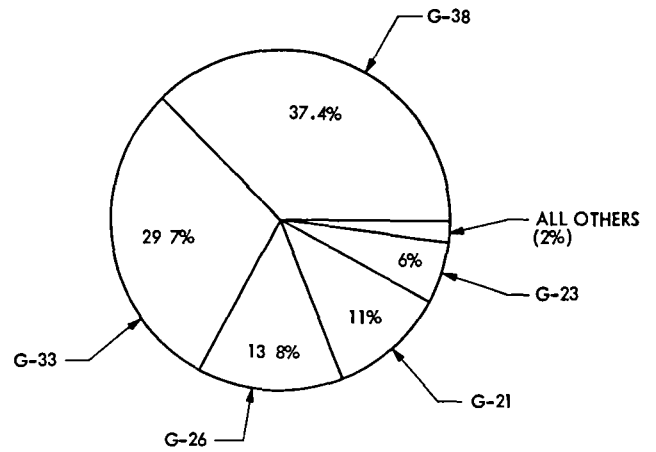


Fig. 10. Thermal equipment consumption for buildings at Echo Station

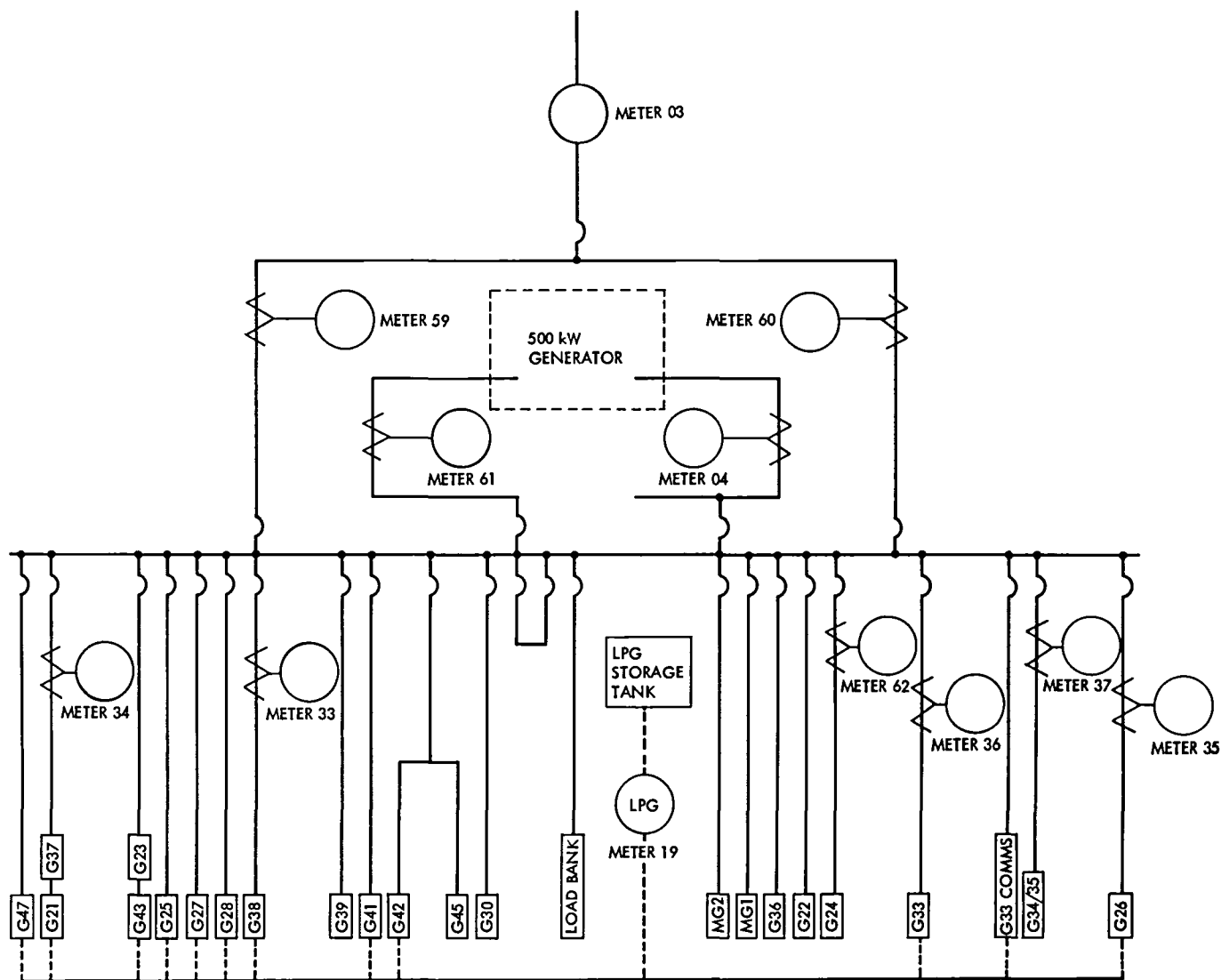


Fig. 11. Metering system for Echo Station

End of Document